

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2003, 12:08:22 ; Search time 14 Seconds.
(without alignments)
918.404 Million cell updates/sec

Title: US-09-914-152-1

Perfect score: 1657

Sequence: 1 MAFPKMLMYICLLVIGALC.....TLDDYQWALENSRGEDCPV 310

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|--------------|--------------------|
| 1 | 1657 | 100.0 | 310 | 1 B3G5_HUMAN | Q9Y2C3 h beta-1,3- |
| 2 | 1592 | 96.1 | 300 | 1 B3G5_GORGO | Q9N293 g beta-1,3- |
| 3 | 1570 | 94.7 | 301 | 1 B3G5_PANPA | Q9N294 p beta-1,3- |
| 4 | 1552 | 93.7 | 297 | 1 B3G5_PANTR | Q9N295 p beta-1,3- |
| 5 | 1206 | 72.8 | 308 | 1 B3G5_MOUSE | Q9J167 m beta-1,3- |
| 6 | 416 | 25.1 | 372 | 1 B3G8_HUMAN | Q9Y2A9 h probable |
| 7 | 360 | 21.7 | 398 | 1 B3G7_HUMAN | Q9Y997 h beta-1,3- |
| 8 | 105.5 | 6.4 | 1049 | 1 VP39_YEAST | Q07468 saccharomyc |
| 9 | 93.5 | 5.6 | 336 | 1 MTHB_HAEPH | P50193 haemophilus |
| 10 | 90.5 | 5.5 | 1314 | 1 SS22_YEAST | P25390 saccharomyc |
| 11 | 89.5 | 5.4 | 786 | 1 TLR1_HUMAN | P15399 saccharomyc |
| 12 | 88.5 | 5.3 | 579 | 1 YHV0_YEAST | P38848 homo sapien |
| 13 | 86.5 | 5.2 | 516 | 1 YNF_AZOCH | P24423 azotobacter |
| 14 | 86.5 | 5.2 | 711 | 1 ETF2_MYXVL | Q9Q8K4 myxoma viru |
| 15 | 85.5 | 5.2 | 632 | 1 Y242_MYCPN | P75440 mycoplasma |
| 16 | 85 | 5.1 | 1073 | 1 HSER_PIG | P55204 sus scrofa |
| 17 | 84.5 | 5.1 | 454 | 1 YF62_HAETN | P45217 haemophilus |
| 18 | 84 | 5.1 | 306 | 1 YF69_SCHPO | Q9P7Q6 schizosacch |
| 19 | 83 | 5.0 | 275 | 1 ISPH_THEMA | Q9X1F7 thermotoga |
| 20 | 82.5 | 5.0 | 525 | 1 DIMH_CAEEL | O17397 caenorhabdi |
| 21 | 82.5 | 5.0 | 713 | 1 PMIP_HUMAN | Q99797 homo sapien |
| 22 | 82 | 4.9 | 576 | 1 HXT4_YEAST | P32467 saccharomyc |
| 23 | 81.5 | 4.9 | 359 | 1 UNG_YEAST | P12887 saccharomyc |
| 24 | 81.5 | 4.9 | 1042 | 1 SVI_BORBU | O51773 borrelia bu |
| 25 | 81 | 4.9 | 290 | 1 LEP_HELPJ | Q9Z1Q5 helicobacte |
| 26 | 81 | 4.9 | 397 | 1 O32A_DROME | P81909 drosophilla |
| 27 | 81 | 4.9 | 399 | 1 TRPB_BUCAP | P42391 buchnera ap |
| 28 | 81 | 4.9 | 1157 | 1 N133_YEAST | P36161 saccharomyc |
| 29 | 80.5 | 4.9 | 290 | 1 LEP_HELPY | O25300 helicobacte |
| 30 | 80.5 | 4.9 | 711 | 1 ETF2_SFVKA | Q9Q8Y2 shope fibro |
| 31 | 80.5 | 4.9 | 386 | 1 POLG_DEN4 | P09866 d genome po |
| 32 | 79.5 | 4.8 | 612 | 1 B1R3_MOUSE | Q62210 mus musculu |
| 33 | 79.5 | 4.8 | 691 | 1 Y104_YEAST | P40460 saccharomyc |

ALIGNMENTS

RESULT 1

B3G5_HUMAN

ID B3G5_HUMAN STANDARD; PRT; 310 AA.

AC Q9Y2C3; Q9NY96; Q9P1X7; Q9P1X6;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Beta-1,3-galactosyltransferase 5 (EC 2.4.1.1-) (Beta-1,3-GalTase 5)

DE (Beta3Gal-T5) (b3Gal-T5) (UDP-galactose:beta-N-acetylglucosamine beta-

DE 1,3-galactosyltransferase 5) (UDP-Gal:beta-GlcNAc beta-1,3-

DE galactosyltransferase 5) (Beta-3-Gx-T5).

GN B3GALT5

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Colorectal adenocarcinoma;

RA Kubota T., Togayachi A., Kudo T., Nishihara S., Watanabe M.,

RA Narimatsu H.;

RT "Cloning, expression, and characterization of a novel UDP-

RT galactose:beta-N-acetylglucosamine beta1,3-galactosyltransferase

RT (beta3Gal-T5) responsible for synthesis of type 1 chain in colorectal

RT and pancreatic epithelia and tumor cells derived therefrom.";

RL J. Biol. Chem. 274:12499-12507(1999).

RN [2]

RP SEQUENCE FROM N.A.

RC MEDLINE=99337698; PubMed=10406968;

RA Zhou D., Berger E.G., Hennet T.;

RT "Molecular cloning of a human UDP-galactose:GlcNAc beta1,3GalNAc

RT beta1,3 galactosyltransferase gene encoding an O-linked

RT core3-elongation enzyme.";

RL Eur. J. Biochem. 263:571-576(1999).

RN [3]

RP SEQUENCE FROM N.A.

RC MEDLINE=20289799; PubMed=10830953;

RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,

RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,

RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,

RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,

RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,

RA Rosenthal A., Kudo J., Shibuya K., Kawasaki K., Asakawa S.,

RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,

RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,

RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,

RA Ramser J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,

RA Wehmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,

p57843 pasteurilla
p58612 anabaena sp
Q12955 homo sapien
O85065 buchnera ap
O59025 methanococc
Q13490 homo sapien
Q09328 homo sapien
O51498 borrelia bu
O57841 methanococc
O8xm18 clostridium
Q01932 rattus norv
p57482 buchnera ap

Qy 1 MAFPKRLMYICLLVLGALCLYFSMYSLNPFKEQSFVYKKDGNFLKLPDTCRQTPPFLV 60
Db 1 MAFPKRLMYICLLVLGALCLYFSMYSLNPFKEQSFVYKKDGNFLKLPDTCRQTPPFLV 60
Qy 61 LLVTSSSHKOLAEIRAIOTWGMKRMVKGKLTFFLLGTTSSAETKEVDQESORHGDII 120
Db 61 LLVTSSSHKOLAEIRAIOTWGMKRMVKGKLTFFLLGTTSSAETKEVDQESORHGDII 120
Qy 121 QKDFLDVYNNLTAKTMGIEWHVHRCFQAAAFVYKMTDSOMFINVDYLTTELLKKNRTTRFF 180
Db 121 QKDFLDVYNNLTAKTMGIEWHVHRCFQAAAFVYKMTDSOMFINVDYLTTELLKKNRTTRFF 180
Qy 181 TGFLKLNREPIROPFSKFWKSEYPMWDRYPFCSTGTYVFGSDVASQVYVYIK 240
Db 181 TGFLKLNREPIROPFSKFWKSEYPMWDRYPFCSTGTYVFGSDVASQVYVYIK 240
Qy 241 LEDVFVGLCLERLNIRLEELHSOPTFFPGGLRFSVCLFRRIIVACHFIKPRTLDDYWOALE 300
Db 241 LEDVFVGLCLERLNIRLEELHSOPTFFPGGLRFSVCLFRRIIVACHFIKPRTLDDYWOALE 300
RESULT 3
B3G5_PANPA STANDARD; PRT; 301 AA.
AC Q9N294;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-1,3-galactosyltransferase 5 (EC 2.4.1.-) (Beta-1,3-GalTase 5)
DE (Beta3Gal-T5) (b3Gal-T5) (UDP-galactose:beta-N-acetylgalucosamine beta-
DE 1,3-galactosyltransferase 5) (UDP-Gal:beta-GlcNAc beta-1,3-
DE galactosyltransferase 5) (Beta-3-Gx-T5) (Fragment).
GN B3GAL1T5.
OS Pan paniscus (Pygmy chimpanzee) (Bonobo).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9597;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu Y., Saitou N.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CATALYZES THE TRANSFER OF GAL TO GLCNAC-BASED ACCEPTORS
CC WITH A PREFERENCE FOR THE CORE3 O-LINKED GLYCAN
CC GLCNAC(BETA1,3)GALNAC STRUCTURE. CAN USE GLYCOPOLIPID LC3CER AS AN
CC EFFICIENT ACCEPTOR (BY SIMILARITY).
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (Potential).
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 31.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB041415; BAA94500.1;
CC InterPro; IPR002659; Galactosyl_T.
CC Pfam; PF01762; Galactosyl_T.1.
CC Transferrase; Glycosyltransferase; Glycoprotein; Transmembrane;
CC Signal-anchor; Golgi stack; Multigene family.
CC CYTOPLASMIC (POTENTIAL).
CC DOMAIN 1 7
CC TRANSMEM 8 28
CC SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC (POTENTIAL).
CC -----
CC DOMAIN 29 >301
CC CARBOHYD 130 130
CC CARBOHYD 174 174
CC CARBOHYD 231 231
CC CARBOHYD 301 301
CC NON_TER 301 301
CC SEQUENCE 301 AA; 35222 MW; E04CC70F41CB29FD CRC64;

Query Match 94.7%; Score 1570; DB 1; Length 301;
Best Local Similarity 98.0%; Pred. No. 7,le-138;
Matches 295; Conservative 2; Mismatches 4; Indels 0; Caps 0;
Qy 1 MAFPKRLMYICLLVLGALCLYFSMYSLNPFKEQSFVYKKDGNFLKLPDTCRQTPPFLV 60
Db 1 MAFPKRLMYICLLVLGALCLYFSMYSLNPFKEQSFVYKKDGNFLKLPDTCRQTPPFLV 60
Qy 61 LLVTSSSHKOLAEIRAIOTWGMKRMVKGKLTFFLLGTTSSAETKEVDQESORHGDII 120
Db 61 LLVTSSSHKOLAEIRAIOTWGMKRMVKGKLTFFLLGTTSSAETKEVDQESORHGDII 120
Qy 121 QKDFLDVYNNLTAKTMGIEWHVHRCFQAAAFVYKMTDSOMFINVDYLTTELLKKNRTTRFF 180
Db 121 QKDFLDVYNNLTAKTMGIEWHVHRCFQAAAFVYKMTDSOMFINVDYLTTELLKKNRTTRFF 180
Qy 181 TGFLKLNREPIROPFSKFWKSEYPMWDRYPFCSTGTYVFGSDVASQVYVYIK 240
Db 181 TGFLKLNREPIROPFSKFWKSEYPMWDRYPFCSTGTYVFGSDVASQVYVYIK 240
Qy 241 LEDVFVGLCLERLNIRLEELHSOPTFFPGGLRFSVCLFRRIIVACHFIKPRTLDDYWOALE 300
Db 241 LEDVFVGLCLERLNIRLEELHSOPTFFPGGLRFSVCLFRRIIVACHFIKPRTLDDYWOALE 300
Qy 301 N 301
Db 301 N 301
RESULT 4
B3G5_PANTR STANDARD; PRT; 297 AA.
AC Q9N295;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-1,3-galactosyltransferase 5 (EC 2.4.1.-) (Beta-1,3-GalTase 5)
DE (Beta3Gal-T5) (b3Gal-T5) (UDP-galactose:beta-N-acetylgalucosamine beta-
DE 1,3-galactosyltransferase 5) (UDP-Gal:beta-GlcNAc beta-1,3-
DE galactosyltransferase 5) (Beta-3-Gx-T5) (Fragment).
GN B3GAL1T5.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu Y., Saitou N.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CATALYZES THE TRANSFER OF GAL TO GLCNAC-BASED ACCEPTORS
CC WITH A PREFERENCE FOR THE CORE3 O-LINKED GLYCAN
CC GLCNAC(BETA1,3)GALNAC STRUCTURE. CAN USE GLYCOPOLIPID LC3CER AS AN
CC EFFICIENT ACCEPTOR (BY SIMILARITY).
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (Potential).
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 31.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB041414; BAA94499.1;
CC InterPro; IPR002659; Galactosyl_T.
CC Pfam; PF01762; Galactosyl_T.1.
CC Transferrase; Glycosyltransferase; Glycoprotein; Transmembrane;
CC Signal-anchor; Golgi stack; Multigene family.
CC CYTOPLASMIC (POTENTIAL).
CC DOMAIN 1 7
CC TRANSMEM 8 28
CC SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC (POTENTIAL).

```

FT DOMAIN 29 >297 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 297 297
SQ SEQUENCE 297 AA; 34850 MW; BB7963250A837A28 CRC64;

Query Match 93.7%; Score 1552; DB 1; Length 297;
Best Local Similarity 97.6%; Pred. No. 3.2e-136;
Matches 290; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAPPKRLMYICLLVLCALCLYFSMYSLNPFKEQSFYVKKDGNFLKLPDPTDCRQTPPELV 60
DB 1 MAPPKRLMYICLLVLCALCLYFSMYSLNPFKEQSFYVKKDGNFLKLPDPTDCRQTPPELV 60
QY 61 LLVTSKHKQLAERMAIRQTWGMKRMVKGKOLKFTFFLLGTTSSAAETKEVDQESORHGDI 120
DB 61 LLVTSKHKQLAERMAIRQTWGMKRMVKGKOLKFTFFLLGTTSSAAETKEVDQESORHGDI 120
QY 121 OKDFLDVYVNLTKTWGMIEWVHRCFQAAAFVVKTDSDMFINVDTYLTLLKKKNTTRFF 180
DB 121 OKDFLDVYVNLTKTWGMIEWVHRCFQAAAFVVKTDSDMFINVDTYLTLLKKKNTTRFF 180
QY 181 TGFCLKNEPPIROPFKSWFVSKSEYPMWDRYPFCSCGTGYVFGSDVASOVYVSKSVPIK 240
DB 181 TGFCLKNEPPIROPFKSWFVSKSEYPMWDRYPFCSCGTGYVFGSDVASOVYVSKSVPIK 240
QY 241 LEDVFVGLCLERLNIRLEELHSQPTFFPGGLRFSVCLFRFRIVACHFTKPRTLDDYQW 297
DB 241 LEDVFVGLCLERLNIRLEELHSQPTFFPGGLRFSVCLFRFRIVACHFTKPRTLDDYQW 297

RESULT 5
B3G5_MOUSE STANDARD: PRT; 308 AA.
AC Q9J167;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-1,3-galactosyltransferase 5 (EC 2.4.1.-) (Beta-1,3-GalTase 5)
DE (Beta3Gal-T5) (UDP-galactose:beta-N-acetylglucosamine beta-1,3-
DE galactosyltransferase 5) (UDP-Gal:beta-GlcNAc beta-1,3-
DE galactosyltransferase 5) (Beta-3-Gx-T5) (Stage-specific embryonic
DE antigen-3 synthase) (SSEA-3 synthase).
GN B3GALT5 OR B3GT5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ;
RX MEDLINE=20390006; PubMed=10837462;
NA Zhou D., Henion T.R., Jungaiwa F.B., Berger E.G., Hennet T.;
RT "The beta1,3-galactosyltransferase beta3GalT-v is a stage-specific
RT embryonic antigen-3 (SSEA-3) synthase.";
RL J. Biol. Chem. 275:22631-22634(2000).
CC -1- FUNCTION: CATALYZES THE TRANSFER OF GAL TO GLCNAC-BASED ACCEPTORS
CC WITH A PREFERENCE FOR THE CORE3 O-LINKED GLYCAN
CC GLCNAC(BETA1,3)GALNAc STRUCTURE. CAN USE GLYCOLIPID LC3CER AS AN
CC EFFICIENT ACCEPTOR. ALSO CATALYZES THE TRANSFER OF GAL TO THE
CC TERMINAL GALNAc UNIT OF THE GLOBOSIDE GB4, THEREBY SYNTHESIZING
CC THE GLYCOLIPID GB5, ALSO KNOWN AS THE STAGE-SPECIFIC EMBRYONIC
CC ANTIGEN-3 (SSEA-3).
CC -1- PATHWAY: Glycosylation.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (Potential).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND KIDNEY.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 31.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF254738; AAF86241.1;
DR MCD; MGI:2136878; B3galT5.
DR InterPro: IPR002659; Galactosyl_T.
DR Pfam: PF01762; Galactosyl_T; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack; Multigene family.
FT DOMAIN 1 7 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 8 25 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT DOMAIN 26 308 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 128 128 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 229 229 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 308 AA; 35964 MW; 789073A5178825B1 CRC64;

Query Match 72.8%; Score 1206; DB 1; Length 308;
Best Local Similarity 71.7%; Pred. No. 3.4e-104;
Matches 223; Conservative 40; Mismatches 44; Indels 4; Gaps 2;

QY 1 MAPPKRLMYICLLVLCALCLYFSMYSLNPFKEQSFYVKKDGNFLKLPDPTDCRQTPPELV 59
DB 1 MAHMKTRLVYASILLMGALCLYFSDMS---PRELPFFVFKSHGKFLQIPDIDCKQKPPFL 57
QY 60 VLLVTSKHKQLAERMAIRQTWGMKRMVKGKOLKFTFFLLGTTSSAAETKEVDQESORHGDI 119
DB 58 VLLVTSKHKQLAERMAIRKTWGTRETSVGGQOQVTRFTFLGTTSDTEMDATTTLESQHKDI 117
QY 120 OKDFLDVYVNLTKTWGMIEWVHRCFQAAAFVVKTDSDMFINVDTYLTLLKKKNTTRFF 179
DB 118 OKDFKDAYENLTKTWGMIEWVHFCPTATAYVVKTDSDMEFVNVGYLTLLKKKNTTRF 177
QY 180 FTGFLKNEPPIROPFKSWFVSKSEYPMWDRYPFCSCGTGYVFGSDVASOVYVSKSVPI 239
DB 178 FTGVIKPHDPPIROKFNKFWVSKFEYPMWDRYPFCSCGTGYVFGSDVAIQVYVNSVSPFI 237
QY 240 KLEDFVGLCLERLNIRLEELHSQPTFFPGGLRFSVCLFRFRIVACHFTKPRTLDDYQWAL 299
DB 238 KLEDFVGLCLERLNIRPEELHTKQTFPPGGLRFSVCLFRFVCHFKIVACHFMKPODLLTYQWAL 297
QY 300 ENSRGEDCPPV 310
DB 298 ENSKEQDCPAV 308

RESULT 6
B3G8_HUMAN STANDARD: PRT; 372 AA.
AC Q9Y2A9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable beta-1,3-galactosyltransferase 8 (EC 2.4.1.-) (Beta-1,3-
DE GalTase 8) (Beta3Gal-T8) (UDP-galactose:beta-N-
DE acetylglucosamine beta-1,3-galactosyltransferase 8) (UDP-Gal:beta-
DE GlcNAc beta-1,3-galactosyltransferase 8) (Beta-3-Gx-T8).
GN B3GALT8 OR TMEM3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99173880; PubMed=10072769;
RA Yokoyama-Kobayashi M., Yamaguchi T., Sekine S., Kato S.;
RT "Selection of cDNAs encoding putative type II membrane proteins on the
RT cell surface from a human full-length cDNA bank.";
RL Gene 228:161-167(1999).
RN [2]
RP REVIEW.

```



```

Query Match      6.4%; Score 105.5; DB 1; Length 1049;
Best Local Similarity 24.4%; Pred. No. 0.1;
Matches 44; Conservative 32; Mismatches 53; Indels 51; Gaps 9;

QY 37 VYKKDGNFLKL-----PD-----TDCRQTPPPFLVLLVTS 65
    || :||| ||| || |
DB 38 VYSRDGNTKLFOVYPDLLQNAKNDPLPPVIENFYFANELSTIFAQCKET--LILLSTT 94
    || :||| ||| || |

QY 66 S---HQQLAERMAIROTWGKERMVCKQLKTFLLGTTSSAAETKVEDQESQRHGIQ- 121
    : : : : || | | | : : : : : : : : : :
DB 95 NLUHEYDRIIDRGINHCWLFERSHKNKKEBKNTLYIYSTINTAKMRVLYINWEGRTYKNNMEA 154
    : : : : || | | | : : : : : : : : : :

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X85374; CAA59691.1; -
REBASE; 3661; M.HphIB.
InterPro; IPR002294; D12N6_mtfrase.
InterPro; IPR002052; N6_Mtase.
Pfam; PF02086; MethyltransfD12; 1.
PRINTS; PR00505; D12N6MTFRASE.
PROSITE; PS00092; N6_MTASE; 1.
Transference; Methyltransferase; Restriction system.
SEQUENCE 336 AA; 39609 MW; 3BD144C7772C4FDE CRC64;

```
Query Match      5.6%; Score 93.5; DB 1; Length 336;  
Best Local Similarity 22.3%; Pred. No. 0.32;  
Matches 39; Conservative 27; Mismatches 58; Indels 51; Gaps 7;  
  
QY   112 ESORHG-DIIQKDFLDVYNLTLL-----TMMGIEWIRFCQAFAVKMTDSDMI I 161  
         ||::| | | | | | | | | | | | | | | | | | | | | | |  
Db    46 EAKRGYQVIANDILNINQLALVLNNQEILTACDVDFISNPSPKGFMTKMNSDVFF 105  
        : | | | | | | | | | | | | | | | | | | | | | |  
QY   162 -----NDVYLTELLLKKRTTRFTGLFLKNEFFIQQ-PFSKWFP-----KSEYP 206  
        : | | | | | | | | | | | | | | | | | | | | | |  
Db    106 FKRECRELDAIRAILKLNTTYKQALAFALMRRAIKMPYSRFTISWEKVQLRDDEYS 165  
        : | | | | | | | | | | | | | | | | | | | | | |  
QY   207 WDRY-----PPFGSCGVGVESGDVASOVYNNYSKSVPYIKLEDPF 245  
        : | | | | | | | | | | | | | | | | | | | | | |  
Db    166 YSKYGRRRAYHNQSFEFHFRNLNSYNQAVFNNGNIHQAYN-----EDVF 210  
        : | | | | | | | | | | | | | | | | | | | | | |  
  
RESULT 10  
SS22_YEAST  
ID SS22_YEAST STANDARD; PRT; 1314 AA.  
AC P25390;
```


YVNF_AZOCH
ID YVNF_AZOCH STANDARD; PRT; 516 AA.
AC P24423;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein in vnfD 5' region.
OS Azotobacter chroococcum mcd 1.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Azotobacter.
OX NCBI_TaxID=355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90356423; PubMed=2388847;
RA Fallik E., Robson R.L.;
RT *Completed sequence of the region encoding the structural genes for
RT the vanadium nitrogenase of Azotobacter chroococcum*;
RL Nucleic Acids Res. 18:4616-4616(1990).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X51756; CAA36057.1;
DR PIR; S14694; S14694.
KW Nitrogen fixation; Hypothetical protein.
SQ SEQUENCE 516 AA; 58676 MW; ED7AF362ED08A14 CRC64;

Query Match 5.2%; Score 86.5; DB 1; Length 516;
Best Local Similarity 24.8%; Pred No. 2.4;
Matches 66; Conservative 30; Mismatches 97; Indels 73; Gaps 12;

QY 55 TPFFLLVLTSSHKOLAERMAIROTGWKRMVKGKQLKTFLLGLTSSAAETKEVDQ----111
DB 284 TPTFLVDLLT-----ERRAF--TPALEQVWMTALLSAFEVGDISTEALMFQAGYLTI 334
QY 112 -ESQRHGDIQKDFLDVYVNLTKTMGIEWVHRCPOAFAFWKTKDSMFINDVYLTLL 170
DB 335 GOVQRIG--AOPRYRLVPNLEVKASLNEALLRLYPEASVYMTQTGLY-----DLL 385
QY 171 LKKNRTTRFTFGFLKNEFPIROPFSKVFSEKSEYPPDWDPYPCSGTGVVSGDVASQVY 230
DB 386 L-----ANDFPGLOTLTFAFFASIPHDWYRNNDIARYEGY-----YASVY 426
QY 231 NVSKSVVPKLEDFVGLCLERLNRLEELHSQPTFFPGGLRFSVCLFRRIYVACHFIKPR 290
DB 427 S-----HFAAL---GLDVRLEDATSH-----GRIDMTV-LFNGHYLFEFKVY 465
QY 291 TLLDYWQALE-----NSRGE 305
DB 466 ELTPAGALQQLKQKGYADKYRARGE 491

RESULT 14
ETF2_MXXVL
ID ETF2_MXXVL STANDARD; PRT; 711 AA.
AC Q908K4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Early transcription factor 82 kDa subunit (VETF large subunit).
GN M096L.
OS Myxoma virus (strain Lausanne).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Leporipoxvirus.
OX NCBI_TaxID=31530;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20032073; PubMed=10562494;

RA Cameron C., Hota-Mitchell S., Chen L., Barrett J., Cao J.X.,
RA Macaulay C., Miller D., Evans D., McFadden G.;
RT "The complete DNA sequence of myxoma virus";
RL Virology 264:298-318(1999).
CC
CC -!- FUNCTION: ACTS WITH RNA POLYMERASE TO INITIATE TRANSCRIPTION FROM
CC EARLY GENE PROMOTERS. A DNA-DEPENDENT ATPASE ACTIVITY IS
CC ASSOCIATED WITH VETF.
CC
CC -!- SUBUNIT: HETERODIMER OF A 70 kDa AND A 82 kDa SUBUNIT.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AF170726; AAF14984.1;
KW Transcription regulation; Activator.
SQ SEQUENCE 711 AA; 82687 MW; F69EC4F63B087E61 CRC64;

Query Match 5.2%; Score 86.5; DB 1; Length 711;
Best Local Similarity 20.1%; Pred. No. 3.6;
Matches 65; Conservative 55; Mismatches 118; Indels 85; Gaps 16;

QY 29 NPFKEOSFVYKDCNFKLPDTCROTTPPLVLLVTSSTSHKOLAEIRMAIROTGWKRMVKG 88
DB 201 NGSRYKSLLYKDEYFINTLGNKFIILTLDERLNLTVMDS-----DGVVTFSSDGTIMINN 255
QY 89 KOLKTEFLLTSSAAETKEVDQESOR-HGDIQKQFLDYYVNLTKTMGIEWVHRCF 147
DB 256 VKLFTSL-----TDIDQMERIKGDVYIKFLST--PITSRIKLNIE-----T 297
QY 148 QAFAVMTKDSMFINDVYLTLLKKNRT-----RFTGFLKL---NEFPIRQ 193
DB 298 SFIEFETATNILLSADKRISIIILAKNHISIKVKNYIPNIEKYFT-FLVAINSMFNIIQ 356
QY 194 PSKVFSEKSEYPPDWDPYPCSGTGY-----VFGDVASQVYVNVSKVPKLEDFVVG- 247
DB 357 QASDFTKVTYV-WSR---ICQNTKNKRKPKVIVSSLDDEMRKVSDFYKSTKEVFIN 412
QY 248 -----LCLERLN-----IRLEEL-----HSQPTFFPGGLRFSVCLFR 280
DB 413 SGIMFSCDPLKKNYVGVLSFYRLQKMCIPCCLKNQSHTE-----TFSSCVYOK 464
QY 281 IVACHFIKPRITLLDYWQALENSR 303
DB 465 EIASDVNP-YILNFGKVVYTKSK 486

RESULT 15
Y242_MYCPN
ID Y242_MYCPN STANDARD; PRT; 632 AA.
AC P75440;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MG242 homolog (F10_orf632o).
GN MPN338 OR MP498.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL; AE000049; AAB96146.1; -.
KW Hypothetical protein; Transmembrane; Complete proteome.

Query Match

| | | | |
|----|-----|---|-----|
| QY | 3 | FPKRLMYICLLVLGALCYFYSWYSLNPPKBOFSVYVKDGNFLKL-----PDTDCROT | 55 |
| DB | 257 | YATMTNTLLMLIIBELRVYFNS-----KEPALILKLLDNKVSRLKEDPDQNPETDLHE- | 309 |
| QY | 56 | PPPLVLLVTSSSHQLAERMAIRQTWCKERMVKQKQLKTEFLLGTTSSAAETFEVDOESOR | 115 |
| DB | 310 | -----LIFPAEKNYLEKEKTSR-WHKRRVKSLSEEL-----LEEIKOINLETKN | 351 |
| QY | 116 | HG-----DIIQKDPLDVIY-----YNLTILKTMNGI-----EWVHRFCPOAA | 150 |
| DB | 352 | ESLAYPDEIVELELDNVHNVSVTKQVFRHOLDLQTLHGIVINPERYIGMWSNHIFVDWEE | 411 |
| QY | 151 | FV-----MKTDSDMF-----INVDYLT-----ELLKKNTTFFTFQF | 183 |
| DB | 412 | FKDLISQITDAENGSDLYGFEXDKLDESICQVNNKYLTIFSSDSSSFLLIKNDQTKVISNY | 471 |
| QY | 184 | LKLN-EFFIROPFSKWFVSKSEY | 205 |
| DB | 472 | VNAOLYFETR-----RWIINDIEY | 490 |

Search completed: April 14, 2003, 12:12:18
Job time : 18 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2003, 12:08:21 ; Search time 28 Seconds
(without alignments)
325.754 Million cell updates/sec

Title: us-09-914-152-1

Perfect score: 1657

Sequence: 1 MAFPMRLMYICLLVLGALC.....TLLDYQWALENSRGDCPPV 310

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents-AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 589 | 35.5 | 326 | 2 | US-09-055-097-6 |
| 2 | 388 | 23.4 | 378 | 2 | US-09-055-097-1 |
| 3 | 359.5 | 21.7 | 378 | 4 | US-09-482-180A-2 |
| 4 | 350 | 21.1 | 325 | 2 | US-09-055-097-5 |
| 5 | 314 | 18.9 | 397 | 4 | US-09-459-133-2 |
| 6 | 307.5 | 18.6 | 389 | 4 | US-09-459-133-13 |
| 7 | 138 | 8.3 | 372 | 1 | US-08-207-904-10 |
| 8 | 81.5 | 4.9 | 1513 | 5 | PCT-US93-03076-2 |
| 9 | 79.5 | 4.8 | 612 | 3 | US-09-212-971-14 |
| 10 | 79.5 | 4.8 | 612 | 4 | US-08-800-929A-14 |
| 11 | 79.5 | 4.8 | 612 | 4 | US-08-569-749-14 |
| 12 | 79.5 | 4.8 | 612 | 4 | US-09-617-053A-14 |
| 13 | 79.5 | 4.8 | 612 | 5 | PCT-US96-12860-14 |
| 14 | 79 | 4.8 | 740 | 3 | US-08-276-968A-18 |
| 15 | 78 | 4.7 | 438 | 5 | PCT-US95-05922A-2 |
| 16 | 78 | 4.7 | 618 | 4 | US-08-569-749-2 |
| 17 | 78 | 4.7 | 618 | 4 | US-09-069-023-29 |
| 18 | 78 | 4.7 | 618 | 5 | PCT-US96-12860-2 |
| 19 | 78 | 4.7 | 741 | 3 | US-08-276-968A-20 |
| 20 | 77 | 4.6 | 454 | 3 | US-08-276-968A-22 |
| 21 | 76 | 4.6 | 740 | 1 | US-08-016-863-16 |
| 22 | 76 | 4.6 | 740 | 3 | US-08-276-968A-16 |
| 23 | 76 | 4.6 | 923 | 4 | US-09-397-885-1 |
| 24 | 75.5 | 4.6 | 331 | 4 | US-09-134-001C-3626 |
| 25 | 75.5 | 4.6 | 821 | 1 | US-09-377-465A-2 |
| 26 | 75.5 | 4.6 | 3177 | 2 | US-08-477-451-4 |
| 27 | 73.5 | 4.4 | 329 | 4 | US-09-651-200-19 |

Query Match 35.5% Score 589; DB 2; Length 326;

ALIGNMENTS

RESULT 1

US-09-055-097-6
; Sequence 6, Application US/09055097
; Patent No. 5955282

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Guegler, Karl J.

APPLICANT: Corley, Neil C.

APPLICANT: Shah, Purvi

APPLICANT: Patterson, Chandra

TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: California

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

CURRENT APPLICATION DATA: Windows Version 2.0

APPLICATION NUMBER: US/09/055,097

FILING DATE: Filed Herewith

CLASSIFICATION:

Prior APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Cerrone, Michael C.

REGISTRATION NUMBER: 39,132

REFERENCE/DOCKET NUMBER: PF-0490 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 326 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 2745735

US-09-055-097-6

Sequence 2, Appli
Sequence 4, Appli
Sequence 6, Appli
Sequence 24, Appli
Sequence 120, App
Sequence 1, Appli
Sequence 4, Appli
Sequence 2, Appli
Sequence 7, Appli
Sequence 4, Appli
Sequence 2, Appli
Sequence 7, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 4, Appli

Best Local Similarity 38.6%; Pred. No. 2.3e-62;
Matches 127; Conservative 52; Mismatches 108; Indels 42; Gaps 7;

QY 12 CLVLGALCLYFSMYSLN-----PKKQSFVYKK--DGN-----FL 45
Db 7 CLYLVSWASALWLSITRPTSSYTGSKPFSHLTVARKNFTFGNIRPINPHSFEFL 66
QY 46 KLPDTCRPTPFLVLLVTSSSHQKLAERMAIROTCKERWVGKOLKTFELGTTSSAAE 105
Db 67 INEPNCKENIFLVILISTTHKEFDARAIRETWGDENNFGIKIATLFLGKNADPVL 126
QY 106 TKEVDQESQRHGDIIQOKDFLDVYVNLTKTMGIEWVHRFCPOAAAFVMTKDSDFINVDY 165
Db 127 NQWVEQESQIFHDIIVEDFIDSVHNLTKTLGMRVVAFTCSKAKVMTKDSDFINVDN 186
QY 166 LTELKLLKRR--TTTFFTCFLKNEPIRQPTSKVFWVSKSEYPMWDPYPCSGTGVVFSG 223
Db 187 LTYLKLKPTKPRRYFTGYV-INGGPIRDVRSKMYMPRDLYPDNSNYPFCSGTGYIFA 245
QY 224 DVASOVYVNSKSVPIKLEDFVGLCLERLNTLRLELHQSPTFFPGG-----LRFVSCVLF 278
Db 246 DVAELIYKTSLTRLLHLEDVYVGLCLRLKLGTH-----PFQNSGFNNHMKMAYSLCRY 297
QY 279 RRIVACHFIKPRTLDDYWOALNSRGEDC 307
Db 298 RRVITVHQISPEEMIRIWNMSSKKHLRC 326

RESULT 2
US-09-055-097-1
; Sequence 1, Application US/09055097
; Patent No. 5955282
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; APPLICANT: Patterson, Chandra
; TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/055.097
; FILING DATE: Filed Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0490 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 378 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:

LIBRARY: DUODNOT02
CLONE: 1705085
US-09-055-097-1

Query Match 23.4%; Score 388; DB 2; Length 378;
Best Local Similarity 29.0%; Pred. No. 4.9e-38;
Matches 104; Conservative 60; Mismatches 129; Indels 66; Gaps 10;

QY 7 RLMIYICLV-----LGALCLYFSMYSLNPKKEQSFVYKKDGNFLKLP-----D 49
Db 8 RLLAALLIVVMTLFGPSGLGELLSSLASLLP-----APASPCPPLALPRLLPNQ 61
QY 50 TDCR--QTPEFVLLVTSSSHQKLAERMAIROTCKERWVGKOLKTFELGTTSS----- 102
Db 62 EACSGCAPPPFLILVCTAPENLNORNAIRASWGLREARGLRVQTLFLLLGEPNAAHPVW 121
QY 103 AATKEVDQESQRHGDIIQOKDFLDVYVNLTKTMGIEWVHRFCPOAAAFVMTKDSDFIN 162
Db 122 GSGOSDLASESAQAQDILQNAFODSYRNLTKTSLGSLNWAEEKHCPMAYVYLLKTDVVYN 181
QY 163 V-DYLTTELLKLNKRTTRF-----FTQFLKLNKNEFPIRQPF 195
Db 182 VPELVSELVLRGCRWQWQWERTEPQEAEOEGGVHLHSEEVPLLYLCRVHVRVNPSTPG 241
QY 196 SKWFSKSEYP--WDYRPPFCSTGTVFSGDVASQVYVNSKSVYIKLEDFVGLCLERL 253
Db 242 GRRVSEEQPHWTGPPPPYASGTGYLSASAVQLILKVASRAPLLPLEDFVGVVSARRG 301
QY 254 NIRLEELHQSPTFFPGGLRFSV---CLFRIVACHFIKPRTLDDYWOALNSRGEDC 309
Db 302 GL-----APTQCVKLAGATHYPLDRCCYKFLITSHRLDPKMQEAMKLVGSGDGERTAP 356

RESULT 3
US-09-482-180A-2
; Sequence 2, Application US/09482180A
; Patent No. 6361985
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Yamamoto, Gayle
; APPLICANT: Gao, Zeren
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Jaspers, Stephen
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING
; TITLE OF INVENTION: BETA-1,3-GALACTOSYLTRANSFERASE HOMOLOG, ZN5SP6
; FILE REFERENCE: 98-80
; CURRENT APPLICATION NUMBER: US/09/482,180A
; CURRENT FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/115,721
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-482-180A-2

Query Match 21.7%; Score 359.5; DB 4; Length 378;
Best Local Similarity 32.3%; Pred. No. 1.3e-34;
Matches 92; Conservative 45; Mismatches 137; Indels 11; Gaps 6;

QY 30 PKKQSFV-YKKDGNFLKLPDTCRPTPFLVLLVTSSSHQKLAERMAIROTCK-ERMVK 87
Db 90 PSRHLFLTYRHCNFPNPSILPEPSGCKOTFLLAIAKSPQCHVERRAAIRSTWGRVGVAR 149
QY 88 GKQKTFELGTTSSNAETKQVDSQRHGDIIQOKDFLDVYVNLTKTMGIEWVHRFCP 147
Db 150 GQKLKLVLLGVAGSAPQAQLLAYESREFDDILQWDFEDFNFLTKELHQLQWVAACP 209
QY 148 QAAVFMKTDSDMFINVDYLTTELLKLNKRTTRFTTFLKLNKNEPIRQPFKSWVSKSEYPW 207
Db 210 QAHFMLKGGDDVVFVHPVNVLEFLDGWDPAQDLLVGDVIRQALPNRNTKVYFIPPSMYRA 269

```

QY 208 DRTPPCSGTGYVFSQVTVNKSVPYIKLEDVYFVGLCLERLNIRLEEHSQPTFF 267
    III: | III: | : : : : III: | III: | : |
DB 270 THYPYAGGGGYMSRATVRLQAIMEDABELLSDIDVFGMCURRLG--LSPMH-HAGFK 326
    III: | III: | : : : : III: | III: | : |
QY 268 PGLR-----FSVCLPRRIVACHFIKPRTLDDYQWALENSRGDC 307
    III: | III: | : : : : III: | III: | : |
DB 327 TFGIRPLDPLDCLYRGLLLVIRLSPLEMWTW-ALVTDEGLKC 370
    III: | III: | : : : : III: | III: | : |

RESULT 4
US-09-055-097-5
: Sequence 5, Application US/09055097
: Patent No. 5955282
: GENERAL INFORMATION:
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Guegler, Karl J.
: APPLICANT: Corley, Neil C.
: APPLICANT: Shah, Purvi
: APPLICANT: Patterson, Chandra
: TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: California
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/055,097
: FILING DATE: Filed Herewith
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Cerrone, Michael C.
: REGISTRATION NUMBER: 39,132
: REFERENCE/DOCKET NUMBER: PF-0490 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (650) 855-0555
: TELEFAX: (650) 845-4166
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 325 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: GenBank
: CLONE: 1150971
: US-09-055-097-5

```

```

Query Match      21.1% Score 350; DB 2; Length 325;
Best Local Similarity 29.0%; Pred. No. 1.5e-33;
Matches 92; Conservative 56; Mismatches 135; Indels 34; Gaps 10;

QY 7 RMYICLLVLGAL-----C-LYFSMYSLNPFKEOSFYVKKDGN-----FLK 46
   : : | | | | : : : : | : | : | : : : : : : : : : : : : : : :
Db 7 KLLRLCILLPLILLVDYCGLLTHLHELN--FERHFHYPLNDGDTSGSASSGLDKFAYLR 64
   : : | | | | : : : : | : | : | : : : : : : : : : : : : : : :

QY 47 LPDTCR--QTPFYLVLVTSHHKQLAERMARQTGKERMVWGKOLKTFFLGLGTSSA 103
   : : | | | | : : : : | : | : | : : : : : : : : : : : : : : :
Db 65 VPSFTAEPVDPQARLTMLIKSAVGNRRREARRRTWGEGRFSDVHLRRVFLLTGTAEDS 124
   : : | | | | : : : : | : | : | : : : : : : : : : : : : : : :

QY 104 AETKEVDQESORHDIIOKDFLDVYYNLTKTMGIEWHFRCQPAAFVKKTDSDMINV 163
   : : | | | | : : | : | | | | | : : : : : : : : : : : : : : :

```

```

Db 125 --EKDVANESREHGDIILQADTDYAFNTLTKMLGMRWASEQNRSEFYLFVDDDDYVSA 182
QY 164 DYLTELL---LKNKRTTRFFTGFLKLNFEPIROPFSKWFVSKSEYPMDRYPPEFCSGTGYV 220
Db 183 KNLVKFLGRGRQSHQPELLFAGHV-FOTSPLRHKFKSNWYVSELEYFDRWPPVVTAGAFI 241
QY 221 FSGDVASQVYVNWKSVPYIKLEDFVGLGRCLERLNIRLEELHUSOPTFPGGLRFSVCLFRR 280
Db 242 LSQALRQLYAASVHLPLFRDFDDVYLGIVALKAGISLQ--HCCDFRHRPAYKGPDSYSS 299
QY 281 IVACH-FIKPRTLDDYW 296
Db 300 VIASHEFGDPEEMTRVW 316

RESULT 5
US-09-459-133-2
; Sequence 2, Application US/09459133
; Patent No. 6416988
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Yamamoto, Gayle
; APPLICANT: Jaspers, Stephen R.
; APPLICANT: Gao, Zeren
; TITLE OF INVENTION: BETA-1,3-GALACTOSYLTRANSFERASE HOMOLOGS
; FILE REFERENCE: 98-77
; CURRENT APPLICATION NUMBER: US/09/459,133
; CURRENT FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/111,697
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (137)...(137)
; OTHER INFORMATION: Xaa is Gly or Ser
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(397)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-459-133-2

Query Match 18.9%; Score 314; DB 4; Length 397;
Best Local Similarity 29.9%; Pred. No. 4.5e-29;
Matches 78; Conservative 50; Mismatches 117; Indels 16; Gaps

```

[illegible]

```
; Sequence 13, Application US/09459133
; Patent No. 6416988
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Yamamoto, Gayle
; APPLICANT: Jaspers, Stephen R.
; APPLICANT: Gao, Zeren
; TITLE OF INVENTION: BETA-1,3-GALACTOSYLTRANSFERASE HOMOLOGS
; FILE REFERENCE: 98-77
; CURRENT APPLICATION NUMBER: US/09/459,133
; CURRENT FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/111,697
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-459-133-13

Query Match      18.6%; Score 307.5; DB 4; Length 389;
Best Local Similarity 29.6%; Pred. No. 2.6e-28;
Matches 86; Conservative 45; Mismatches 135; Indels 25; Gaps 8;

QY 15 VLGCALCYFWSYSLNPFKEQSFYVKKDGNFLKLPDTCR-QTPPELVLLVTSKHKQLAER 73
DB 111 LLSNACRSFPLW--PAGESS-----PVASCSDKDVPIYLLAVKSEPGHFAAR 156

QY 74 MATROTGKERMVKGKOLKTFLLGTTSSAAE---TKVEDQESORHGDIIOKDFLDVYVN 130
DB 157 QAVRETWGSP--VAGTRL--LFLLGSLPGMGDPDLRSVLTWESRRYGDLLLDWDFLDVYN 212

QY 131 LTLKTMGIEWVHRECPQAAFMKTDSDMEINVDYLTELL--LKKNRTTFFTGFLKLINE 188
DB 213 RTUKDLLLLTWLSHCDDVNFVQVQDDAFVHPALLLEHLQTLPTWARSLYLGEIFTQA 272

QY 189 FPIRQPFKSWFVSKSEYPMWRYPFCSTGYVFGSDVASQVYVSKVPYIKLEDVFGVL 248
DB 273 KPLRKPGPFYVPTFEED-YPAYASGGYVIGSLAPLWLLQNAARVAFPPDDVDTGTF 331

QY 249 CLERLNLRLBELNSQPTFFPGGRFVSLFRIRVACHFIKPRTLDDYQAL 299
DB 332 CFRALGLAPRAHPGFLTAWPAERTDPCAVRGLLLVHPVSPQDTIWLWRHL 382

RESULT 7
US-08-207-904-10
; Sequence 10, Application US/08207904
; Patent No. 5477002
; GENERAL INFORMATION:
; APPLICANT: Tuttle, AnnMarie
; APPLICANT: Crossland, Lyle D.
; TITLE OF INVENTION: Antier-Specific cDNA Sequences, Genomic
; NUMBER OF INVENTION: DNA Sequences and Recombinant DNA Sequences
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/207,904
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION: 800

; Sequence 13, Application US/09459133
; Patent No. 6416988
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Yamamoto, Gayle
; APPLICANT: Jaspers, Stephen R.
; APPLICANT: Gao, Zeren
; TITLE OF INVENTION: BETA-1,3-GALACTOSYLTRANSFERASE HOMOLOGS
; FILE REFERENCE: 98-77
; CURRENT APPLICATION NUMBER: US/09/459,133
; CURRENT FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/111,697
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-459-133-13

Query Match      8.3%; Score 138; DB 1; Length 372;
Best Local Similarity 23.7%; Pred. No. 6.6e-08;
Matches 53; Conservative 41; Mismatches 86; Indels 44; Gaps 9;

QY 58 FLVLLVTSKHKQLAERMAIROTWGKERMVKGKOLKTF-----FLLG--TTSSAAET 106
DB 103 FIVVGINTAFSSRRKRDLSRETW----MPKGDKLRLKEKICIVIRFVIGHSATRGCVLD 158

QY 107 KEVDQESORHGDIIOKDFLDVYVNLTLKTMGIE-----WVHRFCPCQAAFMKTDSDMFI 161
DB 159 RAIDSDAQYKDFLRDLHVEGHVHSLTSLVFSKAVSIW-----DADFVYKVVDDVHL 212

QY 162 NVDTLLELLAKNRTTFFTGFLK---LNEPFIQPFKSWFVSKSEYPMW-----DRYP 211
DB 213 NLGMLANTLAKYKSPRVIIIGCMKSGPVLSONGVR-----YYEPEY-WKFGREGNKYF 264

QY 212 PFCSTGYVFGSDVASQVYVSKVPYIKLEDVFGVLCLERLNI 255
DB 265 RHATGQIVIGSKDLASYISINSGLIHRVANEVDVSLGSLWLGLEV 308

RESULT 8
PCT-US93-03076-2
; Sequence 2, Application PC/TUS9303076
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; TITLE OF INVENTION: GAP-Associated Protein p190 and
; TITLE OF INVENTION: Transduction
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: 2 Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/03076
; FILING DATE: 19930331
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WH192-03A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9340
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1513 amino acids
; TYPE: AMINO ACID
```

TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-03076-2

Query Match 4.9%; Score 81.5; DB 5; Length 1513;
Best Local Similarity 23.6%; Pred. No. 3.9;
Matches 37; Conservative 22; Mismatches 63; Indels 35; Gaps 5;

Qy 72 ERMAIRQTWGERMVKGKQLKTTFFLLGTSSAAETKEVDQESORHGDIIQKDFLDVYYNL 131
Db 137 DQGLGQDFEQKQPDGKLLVDLFGDIDVSRGMNRFDD-----QLKfVSNLYNQ 187
Qy 132 TLKTMGI-----EKVHRECPQA-----AFVYKTDSDMFIVNDY-----LTLL 170
Db 188 LAKTKPIVILVTKDCEGVERTIRDAHTFALSCKNLQVETARSNVNDLAFVTLVQLI 247
Qy 171 LKKNRTFRFTGFLKLNFPRIQPFQSKWFSKSEYPW 207
Db 248 DKSRGKTIIPFEAL-----KQSQQIATAKDYEW 279

RESULT 9
US-09-212-971-14
Sequence 14, Application US/09212971B
Patent No. 6107041

GENERAL INFORMATION:
APPLICANT: Korneluk, Robert G
APPLICANT: Mackenzie, Alexander E
APPLICANT: Liston, Peter
APPLICANT: Baird, Stephen
APPLICANT: Tsang, Benjamin K
APPLICANT: Pratt, Christine

TITLE OF INVENTION: DETECTION AND MODULATION OF IAPS AND
TITLE OF INVENTION: NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE
FILE REFERENCE: 07891/009002
CURRENT APPLICATION NUMBER: US/09/212,971B
CURRENT FILING DATE: 1998-12-16
EARLIER APPLICATION NUMBER: 60/017,354
EARLIER FILING DATE: 1996-04-26
EARLIER APPLICATION NUMBER: 60/030,590
EARLIER FILING DATE: 1996-11-14
EARLIER APPLICATION NUMBER: 08/800,929
EARLIER FILING DATE: 1997-02-13
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 612
TYPE: PRT
ORGANISM: Mus musculus
US-09-212-971-14

Query Match 4.8%; Score 79.5; DB 3; Length 612;
Best Local Similarity 19.9%; Pred. No. 1.6;
Matches 50; Conservative 35; Mismatches 79; Indels 87; Gaps 10;

Qy 59 LVLLVTSSSHKLQAE-----RMAIRQTWGERMVKGKQLKT-----FFLLGTSSAAETKE 108
Db 382 VMMSTPPVKAALMGFSRSLVROTQVQIATGENYRTVNDIVSVLLNAEDEREEKE 441
Qy 109 VDQESQRHGD--IIQKDFLDVYYNLT-----LKTMMGIEWHRCPCQAFAVYKTDSDMFI 161
Db 442 RQTEEMASGDLILRKRNRMALFOQLTHVLPILDNL-----EASVITKQEHDIIR 491
Qy 162 NV-----DYLTELLKKNRTRFTTGFGLKLNFPRIQPFQSKWFSKSEYPWDYPPF 213
Db 492 OKTOIPLQARELIDTVLVKGNAANIFKNSLK-----EASVITKQEHDIIR 523
Qy 214 CSGTGYVFGDVASQVYN---VSKSVYPIKLEDFVGLCLERLNIRLEELHSQPTFFPGG 270
Db 524 -----EIDSTLVENLFVEKNMYIPTEDV-SGLSLEQLRRLOEERT-----564
Qy 271 LRFSVCLFRRI 281

Db 565 --CKVCMREV 573

RESULT 10
US-08-800-929A-14
Sequence 14, Application US/08800929A
Patent No. 6133437

GENERAL INFORMATION:
APPLICANT: Korneluk, Robert G
APPLICANT: Mackenzie, Alexander E
APPLICANT: Liston, Peter
APPLICANT: Baird, Stephen
APPLICANT: Tsang, Benjamin K
APPLICANT: Pratt, Christine

TITLE OF INVENTION: DETECTION AND MODULATION OF
TITLE OF INVENTION: IAPS AND NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERA
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/800,929A
FILING DATE: 13-FEB-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/030,590
FILING DATE: 14-NOV-1996
APPLICATION NUMBER: 60/017,354
FILING DATE: 26-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bleker-Brady, Kristina
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 07891/009001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 612 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-800-929A-14

Query Match 4.8%; Score 79.5; DB 4; Length 612;
Best Local Similarity 19.9%; Pred. No. 1.6;
Matches 50; Conservative 35; Mismatches 79; Indels 87; Gaps 10;

Qy 59 LVLLVTSSSHKLQAE-----RMAIRQTWGERMVKGKQLKT-----FFLLGTSSAAETKE 108
Db 382 VMMSTPPVKAALMGFSRSLVROTQVQIATGENYRTVNDIVSVLLNAEDEREEKE 441
Qy 109 VDQESQRHGD--IIQKDFLDVYYNLT-----LKTMMGIEWHRCPCQAFAVYKTDSDMFI 161
Db 442 RQTEEMASGDLILRKRNRMALFOQLTHVLPILDNL-----EASVITKQEHDIIR 491
Qy 162 NV-----DYLTELLKKNRTRFTTGFGLKLNFPRIQPFQSKWFSKSEYPWDYPPF 213
Db 492 OKTOIPLQARELIDTVLVKGNAANIFKNSLK-----EASVITKQEHDIIR 523

QY 214 CSCTGYVFGDVASQVYN---VKSVPYIKLEDVFGCLERLNIRLEELHSQPTFFPGG 270
Db 524 -----EIDSTLYENLFVEKNMKYIPTDV-SGLSLEEQRLRQERT----- 564
QY 271 LRFVSVCLFRRI 281
Db 565 --CKVCMDEV 573

RESULT 11
US-08-569-749-14
; Sequence 14, Application US/08569749
; Patent No. 6187557
; GENERAL INFORMATION:
; APPLICANT: Rothe, Mike
; APPLICANT: Gossdel, David V
; TITLE OF INVENTION: INHIBITORS OF APOPTOSIS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/569,749
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Brezner, David J.
; REGISTRATION NUMBER: 24,774
; REFERENCE/DOCKET NUMBER: A-62464/DJB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)781-1989
; TELEFAX: (415)398-3249
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 612 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-569-749-14

Query Match 4.8%; Score 79.5; DB 4; Length 612;
Best Local Similarity 19.9%; Pred. No. 1.6;
Matches 50; Conservative 35; Mismatches 79; Indels 87; Gaps 10;

QY 59 LVLVLTSSHKOLAE----RMAIROTWCKERMVKGKQLKT-----FFLLGTTSSAAETKE 108
Db 382 VVMSTPVVKAALMGFSRSLVROTQVQILATGENYRTVNDIVSVLLNAEDERREEKE 441
QY 109 VDQESQRHGD--IIQKDFLDVYNYLT-----LKTMMGIEWHRCPCQAFAVMKTDSDMFI 161
Db 442 RQTEEMASGDSLIRKNRMALFOQLTHVLPILDNL-----EASVITKQEHDIIR 491
QY 162 NV-----DYLTELLKKNRTRFTFGFLKLNFFPIRQPFVSKSEYPMWDYPPF 213
Db 492 QKTOIPLQARELIDTVLVKGNAANIFKNSLK----- 523
QY 214 CSCTGYVFGDVASQVYN---VKSVPYIKLEDVFGCLERLNIRLEELHSQPTFFPGG 270
Db 524 -----EIDSTLYENLFVEKNMKYIPTDV-SGLSLEEQRLRQERT----- 564
QY 271 LRFVSVCLFRRI 281
Db 565 --CKVCMDEV 573

RESULT 12

US-09-617-053A-14
; Sequence 14, Application US/09617053A
; Patent No. 6300492
; GENERAL INFORMATION:
; APPLICANT: Korneluk, Robert G
; APPLICANT: Mackenzie, Alexander E
; APPLICANT: Liston, Peter
; APPLICANT: Baird, Stephen
; APPLICANT: Tsang, Benjamin K
; APPLICANT: Pratt, Christine
; TITLE OF INVENTION: DETECTION AND MODULATION OF IAPS AND
; TITLE OF INVENTION: NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE
; TITLE OF INVENTION: DISEASE
; FILE REFERENCE: 07891/009003
; CURRENT APPLICATION NUMBER: US/09/617,053A
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/800,929
; PRIOR FILING DATE: 1997-02-13
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 612
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-617-053A-14

Query Match 4.8%; Score 79.5; DB 4; Length 612;
Best Local Similarity 19.9%; Pred. No. 1.6;
Matches 50; Conservative 35; Mismatches 79; Indels 87; Gaps 10;
QY 59 LVLVLTSSHKOLAE----RMAIROTWCKERMVKGKQLKT-----FFLLGTTSSAAETKE 108
Db 382 VVMSTPVVKAALMGFSRSLVROTQVQILATGENYRTVNDIVSVLLNAEDERREEKE 441
QY 109 VDQESQRHGD--IIQKDFLDVYNYLT-----LKTMMGIEWHRCPCQAFAVMKTDSDMFI 161
Db 442 RQTEEMASGDSLIRKNRMALFOQLTHVLPILDNL-----EASVITKQEHDIIR 491
QY 162 NV-----DYLTELLKKNRTRFTFGFLKLNFFPIRQPFVSKSEYPMWDYPPF 213
Db 492 QKTOIPLQARELIDTVLVKGNAANIFKNSLK----- 523
QY 214 CSCTGYVFGDVASQVYN---VKSVPYIKLEDVFGCLERLNIRLEELHSQPTFFPGG 270
Db 524 -----EIDSTLYENLFVEKNMKYIPTDV-SGLSLEEQRLRQERT----- 564
QY 271 LRFVSVCLFRRI 281
Db 565 --CKVCMDEV 573

RESULT 13

PCT-US96-12860-14
; Sequence 14, Application PC/TUS9612860
; GENERAL INFORMATION:
; APPLICANT: TULARIK, INC.
; TITLE OF INVENTION: INHIBITORS OF APOPTOSIS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30


```
;; SOFTWARE: WORD PERFECT 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US95/05922A
;; FILING DATE: 11 MAY 1995
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: FERRARO, GREGORY D.
;; REGISTRATION NUMBER: 36,134
;; REFERENCE/DOCKET NUMBER: 325800-292
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 201-994-1700
;; TELEFAX: 201-994-1744
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 438 AMINO ACIDS
;; TYPE: AMINO ACID
;; STRANDEDNESS:
;; TOPOLOGY: LINEAR
;; MOLECULE TYPE: PROTEIN
PCT-US95-05922A-2
```

```
Query Match      4.78; Score 78; DB 5; Length 438;
Best Local Similarity 19.3%; Pred. No. 1.5;
Matches 52; Conservative 46; Mismatches 88; Indels 84; Gaps 11;

Qy 32 KEQSFVYKKDGNF-----LKLPTDCROT--PPFL-----VLLVTSSHKQL 70
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 160 KGQEFVDEIQGRYPHILLEOLLSTDTTCENADPPIIHFGPGESSEDVAVMNTPVVKA 219
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 71 AE----RMAIRQTGKRMVKGKLTFFLLGTSSAAETKEVDQESQRHGD----- 118
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 220 LEMGFNRDLVAKQTVOSKILTTGENYKTVNDIVSALLNAEDEKREKEKEQAEEMASDDL 279
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 119 IIOKDELVDVYVNL-----LKTMGIEWHRECPQAAAFVVKTDSDMFINV-DYLTELLK 172
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 280 LIRNRKALFOQLTCVLPILDNLKANVINK--OEHDIIKQKQIPIQARELIDITILVK 336
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 173 KNRTTRFTTGFILNEFPPIROPFSKWFVSKSEYPWDRYPPFCSTGYVFGDVASQVYN- 231
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 337 GNAANATFKNCKL-----EIDSTLYKN 358

Qy 232 --VSKSVPIKLEDFVGLCLERLNIRLEE 259
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 359 LFVDKNMKYIPTDV-SGLSLEQLRLQE 387
```

Search completed: April 14, 2003, 12:11:56
Job time : 31 secs

| | | | |
|----|-----|---|-----|
| Qy | 166 | LTLELLKNNR--TTRFTTGFLKLINEFPIROPFSKWFVSKSEYVPMWDRYPPFCSTGTYVFSG | 223 |
| | | ! ! ! ! ! : ! ! ! ! ! : ! ! ! ! ! : ! ! ! ! ! : ! ! ! ! ! : ! ! ! ! ! : ! ! ! ! ! | |
| Db | 187 | LIYLLKPSNKPRRRYGYV--INGGPIRDVRSKWYMPRLYDPSNYPPFCSTGYVFSR | 245 |
| Qy | 224 | DVASQVYNSKVPYIKLEDFVGLCLERINIRLEELHSQPTFFPG-----LRFSCVLF | 278 |
| Db | 246 | DVAELIYKTSLHTRLLHLEDVYVGLCLRLGIH-----PFQNSGFNHMKWATSLCRY | 297 |
| Qy | 279 | RRIVACHFIKPTRLDDYWOALENSRGEDC | 307 |
| Db | 298 | RRVITVHOISPEEMHRIWDMSSKKHLRC | 326 |

```

RESULT 2
US-09-804-006-8
; Sequence 8, Application US/09804006
; Patent No. US20020119517A1
; GENERAL INFORMATION:
; APPLICANT: White, David
; APPLICANT: Zhou, Jianghong
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: LEPTIN INDUCED GENES
; FILE REFERENCE: 07334/126001
; CURRENT APPLICATION NUMBER: US/09/804,006
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/292,228
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: US 60/108,379
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: US 09/150,857
; PRIOR FILING DATE: 1998-09-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-804-006-8

```

RESULT 3
US-09-804-357-11
; Sequence 11, Application US/09804357
; Patent No. US20010024808A1
; GENERAL INFORMATION:

```

; APPLICANT: White, David
; APPLICANT: Zhou, Jianghong
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: LEPTIN INDUCED GENES
; FILE REFERENCE: 07334/109001
; CURRENT APPLICATION NUMBER: US/09/804,357
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/195,896
; PRIOR FILING DATE: 1998-11-19
; PRIOR APPLICATION NUMBER: US 60/108,379
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: US 09/150,857
; PRIOR FILING DATE: 1998-09-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-804-357-11

```

```

RESULT 4
US-09-529-063-25
  : Sequence 25, Application US/09529063
  : Patent No. US20020102542A1
  : GENERAL INFORMATION:
  : APPLICANT: FUKUSHIMA, DAIKICHI
  : APPLICANT: SHIBAYAMA, SHIRO
  : APPLICANT: TADA, HIDEAKI
  : TITLE OF INVENTION: POLYPEPTIDE, CDNA ENCODING THE POLYPEPTIDE, AND USE OF
  : TITLE OF INVENTION: THE BOTH
  : FILE REFERENCE: Q58769
  : CURRENT APPLICATION NUMBER: US/09/529,063
  : CURRENT FILING DATE: 2000-04-07
  : PRIOR APPLICATION NUMBER: PCT/JP98/04514
  : PRIOR FILING DATE: 1998-10-06
  : PRIOR APPLICATION NUMBER: JP 9-274674
  : PRIOR FILING DATE: 1997-10-07
  : NUMBER OF SEQ ID NOS: 117
  : SOFTWARE: PatentIn Ver. 2.1
  : SEQ ID NO 25
  : LENGTH: 422
  : TYPE: PRT
  : ORGANISM: Homo sapiens
US-09-529-063-25

```

RESULT 3
US-09-804-357-11
; Sequence 11, Application US/09804357
; Patent No. US20010024808A1
; GENERAL INFORMATION:

US-09-804-357-9

Query Match 33.7%; Score 558; DB 10; Length 331;
Best Local Similarity 36.7%; Pred. No. 1e-46;
Matches 114; Conservative 58; Mismatches 127; Indels 12; Gaps 3;

QY 1 MAPPKRLMYICLLVLCALCL-YFSMYSLNP-----FKEQSFVYKKGDNFLKLPD 51
DB 13 MSLSLKLWLLLSLFLVWILSLPHYNVIERVNMWYFYEPIYRQDFRFTLRHSN 72
QY 52 CROTPPFLVLLVTSNKKQLAERMAIROTGWCKRMVKGKQLKTFLLGTTSSAAE---TKE 108
DB 73 CSHQNPFLVILVTSRPSDVAKARQAIKRTWCEKKSWMGVEVLTFFLLGQQAEREDKTLALS 132
QY 109 VDQESQRHGDIIQKDFLDVYNTLTMTMGIEVWHRFCPOAAFWKMTDSMFINVDYLTE 168
DB 133 LEDEHVLGYDIIROQDFLDYNNLTAKTMAFRWYMEFCPNKAKYIMKTDVDFINTGNLVK 192
QY 169 LLLKKNRTTFFTGFLKNEFPPIROPFSKWFVSKSEYPMWRYPFCSCGTGVFSGDVASQ 228
DB 193 YLLNLNHSKEFFGYPLIDNYSYRGFFHKNHISYQYEFKVPFPCYCSGLGYIMSGDLVPR 252
QY 229 VYVNSKVPYIKLEDVFGVGLERLNIRLEELHSQPTFFPGGLRFSVCLFRIRIVACHFIK 288
DB 253 VYEMMSHVKPIKFDVYVVICLNLLKVDIHIPEDTNLFFLYRIHLDVQCLRRVIAAHGFS 312
QY 289 PRTLDDYQWAL 299
DB 313 SKEIITFQVM 323

RESULT 8

US-09-804-006-9

; Sequence 9, Application US/09804006
; Patent No. US20020119517A1
; GENERAL INFORMATION:
; APPLICANT: White, David
; APPLICANT: Zhou, Jianghong
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: LEPTIN INDUCED GENES
; FILE REFERENCE: 07334/126001
; CURRENT APPLICATION NUMBER: US/09/804,006
; PRIOR FILING DATE: 2001-03-12
; PRIOR FILING DATE: 1999-04-15
; PRIOR FILING DATE: 1998-10-29
; PRIOR FILING DATE: 1998-09-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-804-006-9

Query Match 33.7%; Score 558; DB 10; Length 331;
Best Local Similarity 36.7%; Pred. No. 1e-46;
Matches 114; Conservative 58; Mismatches 127; Indels 12; Gaps 3;

QY 1 MAPPKRLMYICLLVLCALCL-YFSMYSLNP-----FKEQSFVYKKGDNFLKLPD 51
DB 13 MSLSLKLWLLLSLFLVWILSLPHYNVIERVNMWYFYEPIYRQDFRFTLRHSN 72
QY 52 CROTPPFLVLLVTSNKKQLAERMAIROTGWCKRMVKGKQLKTFLLGTTSSAAE---TKE 108
DB 73 CSHQNPFLVILVTSRPSDVAKARQAIKRTWCEKKSWMGVEVLTFFLLGQQAEREDKTLALS 132
QY 109 VDQESQRHGDIIQKDFLDVYNTLTMTMGIEVWHRFCPOAAFWKMTDSMFINVDYLTE 168
DB 133 LEDEHVLGYDIIROQDFLDYNNLTAKTMAFRWYMEFCPNKAKYIMKTDVDFINTGNLVK 192

QY 169 LLLKKNRTTFFTGFLKNEFPPIROPFSKWFVSKSEYPMWRYPFCSCGTGVFSGDVASQ 228
DB 193 YLLNLNHSKEFFGYPLIDNYSYRGFFHKNHISYQYEFKVPFPCYCSGLGYIMSGDLVPR 252
QY 229 VYVNSKVPYIKLEDVFGVGLERLNIRLEELHSQPTFFPGGLRFSVCLFRIRIVACHFIK 288
DB 253 VYEMMSHVKPIKFDVYVVICLNLLKVDIHIPEDTNLFFLYRIHLDVQCLRRVIAAHGFS 312
QY 289 PRTLDDYQWAL 299
DB 313 SKEIITFQVM 323

RESULT 9

US-09-739-451-2
; Sequence 2, Application US/09739451
; Patent No. US20010024813A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Dendriac and Brainiac-3
; FILE REFERENCE: PP464
; CURRENT APPLICATION NUMBER: US/09/739,451
; PRIOR FILING DATE: 2000-12-19
; PRIOR FILING DATE: 1998-12-17
; PRIOR FILING DATE: 1998-03-12
; PRIOR FILING DATE: 1998-11-17
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 2
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-739-451-2

Query Match 33.5%; Score 555.5; DB 10; Length 319;
Best Local Similarity 36.2%; Pred. No. 1.7e-46;
Matches 112; Conservative 57; Mismatches 125; Indels 15; Gaps 2;

QY 6 MRLMYICLLVLCALCLYFSMYSLNP-----FKEQSFVYKKGDNFLKLPD 53
DB 3 LRSLSLKLWLLLSLFLVWILSLPHYNVIERVNMWYFYEPIYRQDFRFTLRHSNCS 62
QY 54 QTPPFLVLLVTSNKKQLAERMAIROTGWCKRMVKGKQLKTFLLGTTSSAAE---TKEVD 110
DB 63 HQNPFVLVLTSHPSDVAKARQAIKRTWCEKKSWMGVEVLTFFLLGQQAEREDKTLALS 122
QY 111 QESQRHGDIIQKDFLDVYNTLTMTMGIEVWHRFCPOAAFWKMTDSMFINVDYLTE 170
DB 123 DEHLLYDGIIRQDFLDYNNLTAKTMAFRWYMEFCPNKAKYIMKTDVDFINTGNLVKYL 182
QY 171 LKKNRTTFFTGFLKNEFPPIROPFSKWFVSKSEYPMWRYPFCSCGTGVFSGDVASQV 230
DB 183 LNLNHSKEFFGYPLIDNYSYRGFYQKTHISYQYEFKVPFPCYCSGLGYIMSLRVPRIY 242
QY 231 NVKSKVPYIKLEDVFGVGLERLNIRLEELHSQPTFFPGGLRFSVCLFRIRIVACHFIK 290
DB 243 EMGHVVKPIKFDVYVVICLNLLKVDIHIPEDTNLFFLYRIHLDVQCLRRVIAAHGFS 302
QY 291 TLLDYQWAL 299
DB 303 EIITFQVM 311

RESULT 10

US-09-739-451-10
; Sequence 10, Application US/09739451
; Patent No. US20010024813A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Dendriac and Brainiac-3

```
; FILE REFERENCE: PF464
; CURRENT APPLICATION NUMBER: US/09/739,451
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/213,364
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: 60/077,687
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/108,928
; PRIOR FILING DATE: 1998-11-17
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-739-451-10

Query Match      33.5%; Score 555.5; DB 10; Length 320;
Best Local Similarity 36.2%; Pred. No. 1.7e-46;
Matches 112; Conservative 57; Mismatches 125; Indels 15; Gaps 2;

Oy 6 MRLMYICLLVIGALCLYFSMSYSLNP-----FKQSFVYKDKGNFLKLPDTCR 53
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 LRSLLKWSLLLSLSSFFVMVYLSLPHYNNVIERVNNMYEYEPYRQDFHFTLREHSNCS 62

Oy 54 QTPPLVLLVTSSSHKOLAERMAIROTWCKERMVCKQKLTFFLCTTSSAAB---TKEVD 110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 HONPFLVLVTSHPSDVKARQAIRVTWGEKSWMGVEYLVTFLLQGEAEKEDKMLALSLE 122

Oy 111 QESQRHGDIIOKQFLDVYVNIITLTKMGIENWHRCPQAAFMVKTDSDMFINVDVLTLL 170
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 123 DEHLLYGDIIHQDFLDYTNLTITKIMAFRNVTECPNAKVMKTDVDFVITGNLVYL 182

Oy 171 LKKNRTFRFFTGFLKLNFFTRQFSEKWFVSKSEYVMDRPPFCGSGTGYSFSGDVASQVY 230
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 183 LNLNHSKFFTCYPLIDNYSYRGFYQKTHISYQEPYKPVFPYPCSLGYINSRDLVPRIY 242

Oy 231 NVSKSVYKIKEDFVGLCLERLNRLBELHSQPTFFPGGLRFSVCLFRRIVACHFIKPR 290
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 243 EMGHVPIKEDYVIGICLNLLKVNHIHPEDTNLFLYRIHLDVQCLRRIAAGHGFSSK 302

Oy 291 TLLDYWQAL 299
   : : : : :
Db 303 EITFWQVM 311

RESULT 11
US-09-992-598-209
; Sequence 209, Application US/09992598
; Patent No. US20020160384A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
```

| | | |
|--|---------------------------|------------|
| | PRIOR APPLICATION NUMBER: | 60/088958 |
| | PRIOR FILING DATE: | 1998-06-11 |
| | PRIOR APPLICATION NUMBER: | 60/088961 |
| | PRIOR FILING DATE: | 1998-06-11 |
| | PRIOR APPLICATION NUMBER: | 60/088976 |
| | PRIOR FILING DATE: | 1998-06-11 |
| | PRIOR APPLICATION NUMBER: | 60/089105 |
| | PRIOR FILING DATE: | 1998-06-12 |
| | PRIOR APPLICATION NUMBER: | 60/089440 |
| | PRIOR FILING DATE: | 1998-06-16 |
| | PRIOR APPLICATION NUMBER: | 60/089512 |
| | PRIOR FILING DATE: | 1998-06-16 |
| | PRIOR APPLICATION NUMBER: | 60/089514 |
| | PRIOR FILING DATE: | 1998-06-16 |
| | PRIOR APPLICATION NUMBER: | 60/089532 |
| | PRIOR FILING DATE: | 1998-06-17 |
| | PRIOR APPLICATION NUMBER: | 60/089538 |
| | PRIOR FILING DATE: | 1998-06-17 |
| | PRIOR APPLICATION NUMBER: | 60/089598 |
| | PRIOR FILING DATE: | 1998-06-17 |
| | PRIOR APPLICATION NUMBER: | 60/089599 |
| | PRIOR FILING DATE: | 1998-06-17 |
| | PRIOR APPLICATION NUMBER: | 60/089600 |
| | PRIOR FILING DATE: | 1998-06-17 |
| | PRIOR APPLICATION NUMBER: | 60/089653 |
| | PRIOR FILING DATE: | 1998-06-17 |
| | PRIOR APPLICATION NUMBER: | 60/089801 |
| | PRIOR FILING DATE: | 1998-06-18 |
| | PRIOR APPLICATION NUMBER: | 60/089907 |
| | PRIOR FILING DATE: | 1998-06-18 |
| | PRIOR APPLICATION NUMBER: | 60/089908 |
| | PRIOR FILING DATE: | 1998-06-18 |
| | PRIOR APPLICATION NUMBER: | 60/089947 |
| | PRIOR FILING DATE: | 1998-06-19 |
| | PRIOR APPLICATION NUMBER: | 60/089948 |
| | PRIOR FILING DATE: | 1998-06-19 |
| | PRIOR APPLICATION NUMBER: | 60/089952 |
| | PRIOR FILING DATE: | 1998-06-19 |
| | PRIOR APPLICATION NUMBER: | 60/090246 |
| | PRIOR FILING DATE: | 1998-06-22 |
| | PRIOR APPLICATION NUMBER: | 60/090252 |
| | PRIOR FILING DATE: | 1998-06-22 |
| | PRIOR APPLICATION NUMBER: | 60/090254 |
| | PRIOR FILING DATE: | 1998-06-22 |
| | PRIOR APPLICATION NUMBER: | 60/090349 |
| | PRIOR FILING DATE: | 1998-06-23 |
| | PRIOR APPLICATION NUMBER: | 60/090355 |
| | PRIOR FILING DATE: | 1998-06-23 |
| | PRIOR APPLICATION NUMBER: | 60/090429 |
| | PRIOR FILING DATE: | 1998-06-24 |
| | PRIOR APPLICATION NUMBER: | 60/090431 |
| | PRIOR FILING DATE: | 1998-06-24 |
| | PRIOR APPLICATION NUMBER: | 60/090435 |
| | PRIOR FILING DATE: | 1998-06-24 |
| | PRIOR APPLICATION NUMBER: | 60/090444 |
| | PRIOR FILING DATE: | 1998-06-24 |
| | PRIOR APPLICATION NUMBER: | 60/090445 |
| | PRIOR FILING DATE: | 1998-06-24 |
| | PRIOR APPLICATION NUMBER: | 60/090472 |
| | PRIOR FILING DATE: | 1998-06-24 |
| | PRIOR APPLICATION NUMBER: | 60/090535 |
| | PRIOR FILING DATE: | 1998-06-24 |
| | PRIOR APPLICATION NUMBER: | 60/090540 |
| | PRIOR FILING DATE: | 1998-06-24 |
| | PRIOR APPLICATION NUMBER: | 60/090542 |
| | PRIOR FILING DATE: | 1998-06-24 |
| | PRIOR APPLICATION NUMBER: | 60/090557 |
| | PRIOR FILING DATE: | 1998-06-24 |
| | PRIOR APPLICATION NUMBER: | 60/090676 |
| | PRIOR FILING DATE: | 1998-06-25 |
| | PRIOR APPLICATION NUMBER: | 60/090678 |
| | PRIOR FILING DATE: | 1998-06-25 |
| | PRIOR APPLICATION NUMBER: | 60/090690 |

| | |
|----|-------------------------------------|
| 1 | PRIOR FILING DATE: 1998-06-25 |
| 2 | PRIOR APPLICATION NUMBER: 60/090694 |
| 3 | PRIOR FILING DATE: 1998-06-25 |
| 4 | PRIOR APPLICATION NUMBER: 60/090695 |
| 5 | PRIOR FILING DATE: 1998-06-25 |
| 6 | PRIOR APPLICATION NUMBER: 60/090696 |
| 7 | PRIOR FILING DATE: 1998-06-25 |
| 8 | PRIOR APPLICATION NUMBER: 60/090862 |
| 9 | PRIOR FILING DATE: 1998-06-26 |
| 10 | PRIOR APPLICATION NUMBER: 60/090863 |
| 11 | PRIOR FILING DATE: 1998-06-26 |
| 12 | PRIOR APPLICATION NUMBER: 60/091360 |
| 13 | PRIOR FILING DATE: 1998-07-01 |
| 14 | PRIOR APPLICATION NUMBER: 60/091478 |
| 15 | PRIOR FILING DATE: 1998-07-02 |
| 16 | PRIOR APPLICATION NUMBER: 60/091544 |
| 17 | PRIOR FILING DATE: 1998-07-01 |
| 18 | PRIOR APPLICATION NUMBER: 60/091519 |
| 19 | PRIOR FILING DATE: 1998-07-02 |
| 20 | PRIOR APPLICATION NUMBER: 60/091626 |
| 21 | PRIOR FILING DATE: 1998-07-02 |
| 22 | PRIOR APPLICATION NUMBER: 60/091633 |
| 23 | PRIOR FILING DATE: 1998-07-02 |
| 24 | PRIOR APPLICATION NUMBER: 60/091978 |
| 25 | PRIOR FILING DATE: 1998-07-07 |
| 26 | PRIOR APPLICATION NUMBER: 60/091982 |
| 27 | PRIOR FILING DATE: 1998-07-07 |
| 28 | PRIOR APPLICATION NUMBER: 60/092182 |
| 29 | PRIOR FILING DATE: 1998-07-09 |

| | | | | |
|-----------------------|--------|--------------------|-------|-------------|
| Query Match | 33.5%; | Score 555.5; | DB 9; | Length 331; |
| Best Local Similarity | 36.2%; | Pred. No. 1.8e-46; | | |

```

QY 6 MRLMYICLLVIGALCYFSWYSLNP -----FKBQSPVYKKDQNFKLKLPDTCR 53
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 15 LRSKLWSLLLSLLSFEVWYSLPHYNVIERVNMWYFYEYPIYRQDFHTLREHSNCS 74
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 54 QTPPELVLLTSSHKOLAERMAIROTWGMKRMVKGOLKLTFFLLCTTSSAAE---TKEVD 110
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 75 HQNPFVLVLTSPSHSVKARQAI RVTWGEKXSWGVEVLTFLLGQAEKEDKMLALSLE 134
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 111 QESQRHGDIIQKDFLDVYYNLTUKTMGIEWHRCFQAAFNKMTSDMFINVDTLELL 170
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 135 DEHLGYDIIIRQDFLDYNNLTUKTMAPRWVTEFCPNAKYVNMKTDYFINTGNLVKYL 194
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 171 LKXNRTFTFTGFLKLNFPFROPFSKFWVSKSEYPDWRPPPCSGTGVYFGSDVASQVY 230
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 195 LNLNHSKEFTGYPGLINDISYRGFYOKTHISYOEPYFKVFPYPCYSLGYIMSRDLVPRIY 254
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 231 NVKSVVPYKLEDFVFCICLERLNIRLEELHSQPTFFPGCLRSVCLFRRIVACHPIKPR 290
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 255 EMGHVAKPIKFDVYVGICLLNKLVNIHTPEDTNLFYRIHLDVQCLRRVIAAHGFSK 314
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 291 TLLDYMQAL 299
: : : : :
Db 315 EIITFMQVM 323

RESULT 12
US-09-989-293A-209
; Sequence 209, Application US/09989293A
; Patent No. US20020177164A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Denoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.

```

APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zenin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PIC66
CURRENT FILING DATE: 2001-11-20
PRIORITY APPLICATION NUMBER: 60/049787
PRIORITY FILING DATE: 1997-06-16
PRIORITY APPLICATION NUMBER: 60/062250
PRIORITY FILING DATE: 1997-10-17
PRIORITY APPLICATION NUMBER: 60/065186
PRIORITY FILING DATE: 1997-11-12
PRIORITY APPLICATION NUMBER: 60/065311
PRIORITY FILING DATE: 1997-11-13
PRIORITY APPLICATION NUMBER: 60/066770
PRIORITY FILING DATE: 1997-11-24
PRIORITY APPLICATION NUMBER: 60/075945
PRIORITY FILING DATE: 1998-02-25
PRIORITY APPLICATION NUMBER: 60/078910
PRIORITY FILING DATE: 1998-03-20
PRIORITY APPLICATION NUMBER: 60/083322
PRIORITY FILING DATE: 1998-04-28
PRIORITY APPLICATION NUMBER: 60/084600
PRIORITY FILING DATE: 1998-05-07
PRIORITY APPLICATION NUMBER: 60/087106
PRIORITY FILING DATE: 1998-05-28
PRIORITY APPLICATION NUMBER: 60/087607
PRIORITY FILING DATE: 1998-06-02
PRIORITY APPLICATION NUMBER: 60/087609
PRIORITY FILING DATE: 1998-06-02
PRIORITY APPLICATION NUMBER: 60/087759
PRIORITY FILING DATE: 1998-06-02
PRIORITY APPLICATION NUMBER: 60/087827
PRIORITY FILING DATE: 1998-06-03
PRIORITY APPLICATION NUMBER: 60/088021
PRIORITY FILING DATE: 1998-06-04
PRIORITY APPLICATION NUMBER: 60/088025
PRIORITY FILING DATE: 1998-06-04
PRIORITY APPLICATION NUMBER: 60/088026
PRIORITY FILING DATE: 1998-06-04
PRIORITY APPLICATION NUMBER: 60/088028
PRIORITY FILING DATE: 1998-06-04
PRIORITY APPLICATION NUMBER: 60/088029
PRIORITY FILING DATE: 1998-06-04
PRIORITY APPLICATION NUMBER: 60/088030
PRIORITY FILING DATE: 1998-06-04
PRIORITY APPLICATION NUMBER: 60/088033
PRIORITY FILING DATE: 1998-06-04
PRIORITY APPLICATION NUMBER: 60/088326
PRIORITY FILING DATE: 1998-06-04
PRIORITY APPLICATION NUMBER: 60/088167
PRIORITY FILING DATE: 1998-06-05
PRIORITY APPLICATION NUMBER: 60/088202
PRIORITY FILING DATE: 1998-06-05
PRIORITY APPLICATION NUMBER: 60/088212
PRIORITY FILING DATE: 1998-06-05
PRIORITY APPLICATION NUMBER: 60/088217
PRIORITY FILING DATE: 1998-06-05
PRIORITY APPLICATION NUMBER: 60/088555
PRIORITY FILING DATE: 1998-06-09
PRIORITY APPLICATION NUMBER: 60/088734
PRIORITY FILING DATE: 1998-06-10
PRIORITY APPLICATION NUMBER: 60/088738
PRIORITY FILING DATE: 1998-06-10
PRIORITY APPLICATION NUMBER: 60/088742
PRIORITY FILING DATE: 1998-06-10
PRIORITY APPLICATION NUMBER: 60/088810
PRIORITY FILING DATE: 1998-06-10
PRIORITY APPLICATION NUMBER: 60/088824
PRIORITY FILING DATE: 1998-06-10
PRIORITY APPLICATION NUMBER: 60/088826
PRIORITY FILING DATE: 1998-06-10
PRIORITY APPLICATION NUMBER: 60/088858
PRIORITY FILING DATE: 1998-06-11
PRIORITY APPLICATION NUMBER: 60/088861
PRIORITY FILING DATE: 1998-06-11
PRIORITY APPLICATION NUMBER: 60/088876
PRIORITY FILING DATE: 1998-06-11
PRIORITY APPLICATION NUMBER: 60/089105
PRIORITY FILING DATE: 1998-06-12
PRIORITY APPLICATION NUMBER: 60/089440
PRIORITY FILING DATE: 1998-06-16
PRIORITY APPLICATION NUMBER: 60/089512
PRIORITY FILING DATE: 1998-06-16
PRIORITY APPLICATION NUMBER: 60/089514
PRIORITY FILING DATE: 1998-06-16
PRIORITY APPLICATION NUMBER: 60/089532
PRIORITY FILING DATE: 1998-06-17
PRIORITY APPLICATION NUMBER: 60/089538
PRIORITY FILING DATE: 1998-06-17
PRIORITY APPLICATION NUMBER: 60/089598
PRIORITY FILING DATE: 1998-06-17
PRIORITY APPLICATION NUMBER: 60/089599
PRIORITY FILING DATE: 1998-06-17
PRIORITY APPLICATION NUMBER: 60/089600
PRIORITY FILING DATE: 1998-06-17
PRIORITY APPLICATION NUMBER: 60/089653
PRIORITY FILING DATE: 1998-06-17
PRIORITY APPLICATION NUMBER: 60/089801
PRIORITY FILING DATE: 1998-06-18
PRIORITY APPLICATION NUMBER: 60/089907
PRIORITY FILING DATE: 1998-06-18
PRIORITY APPLICATION NUMBER: 60/089908
PRIORITY FILING DATE: 1998-06-18
PRIORITY APPLICATION NUMBER: 60/089947
PRIORITY FILING DATE: 1998-06-19
PRIORITY APPLICATION NUMBER: 60/089948
PRIORITY FILING DATE: 1998-06-19
PRIORITY APPLICATION NUMBER: 60/089952
PRIORITY FILING DATE: 1998-06-19
PRIORITY APPLICATION NUMBER: 60/090246
PRIORITY FILING DATE: 1998-06-22
PRIORITY APPLICATION NUMBER: 60/090252
PRIORITY FILING DATE: 1998-06-22
PRIORITY APPLICATION NUMBER: 60/090254
PRIORITY FILING DATE: 1998-06-22
PRIORITY APPLICATION NUMBER: 60/090349
PRIORITY FILING DATE: 1998-06-23
PRIORITY APPLICATION NUMBER: 60/090355
PRIORITY FILING DATE: 1998-06-23
PRIORITY APPLICATION NUMBER: 60/090429
PRIORITY FILING DATE: 1998-06-24
PRIORITY APPLICATION NUMBER: 60/090431
PRIORITY FILING DATE: 1998-06-24
PRIORITY APPLICATION NUMBER: 60/090435
PRIORITY FILING DATE: 1998-06-24
PRIORITY APPLICATION NUMBER: 60/090444
PRIORITY FILING DATE: 1998-06-24
PRIORITY APPLICATION NUMBER: 60/090445
PRIORITY FILING DATE: 1998-06-24
PRIORITY APPLICATION NUMBER: 60/090472
PRIORITY FILING DATE: 1998-06-24

PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 33.5%; Score 555.5; DB 9; Length 331;
Best Local Similarity 36.2%; Pred. No. 1.8e-46;
Matches 112; Conservative 57; Mismatches 125; Indels 15; Gaps 2;
QY 6 MRLMYICLVGALCLYFSMYSLNP-----FKQSFYKKGNGFLKLPDTCR 53
DB 15 LRSKWSLLLSLSFFVMYLSLPHYNVIERVNMWYFYEYPIYRQDFHTLREHSNCS 74
QY 54 QTPPFLVLLVTSKSHKOLAERMAITQTWGERMVKGKQLKTFLLGTTSSAAE---TKEVD 110
DB 75 HONPFLVILVTSHPSDVKARQAIRVTWGEKSMWGYEVLTFLLGQEAKEKDKMLALSLE 134
QY 111 QESQRHGDIIQKDFLDVYVNLTKTMGIEWHRFCPOAFAVMTKDSDFINVDYLTLL 170
DB 135 DEHLLYGDIIRQDFLDYTNLTKTINAFRWVTEFCPCNAKYVMTDQVDFINTGNLVKYL 194
QY 171 LKKNRTTFFTGFLKNEFFPIOPFSKWFVSKSEYKWDYPPFCGTYGVFSGDVASQVY 230
DB 195 LNLHSEKFFTGPLYDNYSYRGFYQKTHISYQYEPKVPFPYCSGLGYIMSRDLVPRYI 254
QY 231 NVSKSVPIKLEDFVGLCLERLNIRLELHSQTPFPGLRFSVCLFRIRIVACHFIKPR 290
DB 255 EMGHVKPIKFEDVYVYGICLNLLKVINIHIPEDTNLFFLYRIHLDVQCLRRVIAAHGFSK 314
QY 291 TLLDYWQAL 299
DB 315 EIITFWQVM 323

RESULT 13

US-10-063-547-36
Sequence 36, Application US/10063547
Publication No. US20020182638A1
GENERAL INFORMATION:
APPLICANT: Eaton, Dan L.
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3230R1C1
CURRENT APPLICATION NUMBER: US/10/063,547
CURRENT FILING DATE: 2002-05-02
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 170
SEQ ID NO 36
LENGTH: 331
TYPE: PRT
ORGANISM: Homo Sapien
US-10-063-547-36

Query Match 33.5%; Score 555.5; DB 9; Length 331;
Best Local Similarity 36.2%; Pred. No. 1.8e-46;
Matches 112; Conservative 57; Mismatches 125; Indels 15; Gaps 2;
QY 6 MRLMYICLVGALCLYFSMYSLNP-----FKQSFYKKGNGFLKLPDTCR 53
DB 15 LRSKWSLLLSLSFFVMYLSLPHYNVIERVNMWYFYEYPIYRQDFHTLREHSNCS 74
QY 54 QTPPFLVLLVTSKSHKOLAERMAITQTWGERMVKGKQLKTFLLGTTSSAAE---TKEVD 110
DB 75 HONPFLVILVTSHPSDVKARQAIRVTWGEKSMWGYEVLTFLLGQEAKEKDKMLALSLE 134
QY 111 QESQRHGDIIQKDFLDVYVNLTKTMGIEWHRFCPOAFAVMTKDSDFINVDYLTLL 170
DB 135 DEHLLYGDIIRQDFLDYTNLTKTINAFRWVTEFCPCNAKYVMTDQVDFINTGNLVKYL 194
QY 171 LKKNRTTFFTGFLKNEFFPIOPFSKWFVSKSEYKWDYPPFCGTYGVFSGDVASQVY 230
DB 195 LNLHSEKFFTGPLYDNYSYRGFYQKTHISYQYEPKVPFPYCSGLGYIMSRDLVPRYI 254
QY 231 NVSKSVPIKLEDFVGLCLERLNIRLELHSQTPFPGLRFSVCLFRIRIVACHFIKPR 290
DB 255 EMGHVKPIKFEDVYVYGICLNLLKVINIHIPEDTNLFFLYRIHLDVQCLRRVIAAHGFSK 314
QY 291 TLLDYWQAL 299
DB 315 EIITFWQVM 323

RESULT 14
US-09-989-735-209
Sequence 209, Application US/09989735
Publication No. US20020193299A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.

APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PIC61
CURRENT APPLICATION NUMBER: US/09/989,735
PRIOR FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738

PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24

```
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match      33.58;   Score 555.5;   DB 9;   Length 331;
Best Local Similarity 36.28;   Pred. No. 1.8e-46;
Matches 112;   Conservative 57;   Mismatches 125;   Indels 15;   Gaps 2;

Qy 6 MRLMYICLLVGLALCLFYSMSLNP-----FKQSFVYKKGDNFLKLPDTCR 53
Db 15 LRSKLWSLLLSLSSFFVMWYLSLPHYNNVIERVNMWYVEYEPYRQDFHETLREHSNCS 74
Qy 54 QTPFLVLLVTSSSHKQLAERMAIROTWCKERNVKQKLTFFLLGTSSAAB---TKQVD 110
Db 75 HQNPLVLVTSHPSDVKARQAIRTWCKERNVKQKLTFFLLGTSSAAB---TKQVD 134
Qy 111 QESQRHGDIOKDFLDVYVNTLTMTMGIEWVHRFCPOAAFMKTDSDMFVNDVLTLL 170
Db 135 DEHLVYDIIHQDFLDVYVNTLTMTMGIEWVHRFCPOAAFMKTDSDMFVNDVLTLL 194
Qy 171 LKNNRTTRFFGLKLNFFPIRQPFKSWFVSKSEYPPFCSGTGYVFSGVASOVY 230
Db 195 LNLNHESEKFFGYPLIDNYSYRGFYQKTHISYQVPEKVFPPYCSGLGYINRDLVPIY 254
Qy 231 NVSKSVYPIKLEDFVGLCLERLNIRLEELHSQTPFPFGGRFVSCLFRRIVACHFTKPR 290
Db 255 EMGHVVRPIKPFEDVYVYVGLCLERLNIRLEELHSQTPFPFGGRFVSCLFRRIVACHFTKPR 314
Qy 291 TLLDYQWAL 299
Db 315 ELITFWQVM 323
```

RESULT 15

```
US-09-990-444-209
; Sequence 209, Application US/09990444
; Publication No. US20020193300A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C19
; CURRENT APPLICATION NUMBER: US/09/990.444
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
```

| | |
|----|-------------------------------------|
| 1 | PRIOR APPLICATION NUMBER: 60/088161 |
| 2 | PRIOR FILING DATE: 1998-06-05 |
| 3 | PRIOR APPLICATION NUMBER: 60/088202 |
| 4 | PRIOR FILING DATE: 1998-06-05 |
| 5 | PRIOR APPLICATION NUMBER: 60/088212 |
| 6 | PRIOR FILING DATE: 1998-06-05 |
| 7 | PRIOR APPLICATION NUMBER: 60/088217 |
| 8 | PRIOR FILING DATE: 1998-06-05 |
| 9 | PRIOR APPLICATION NUMBER: 60/088655 |
| 10 | PRIOR FILING DATE: 1998-06-09 |
| 11 | PRIOR APPLICATION NUMBER: 60/088734 |
| 12 | PRIOR FILING DATE: 1998-06-10 |
| 13 | PRIOR APPLICATION NUMBER: 60/088738 |
| 14 | PRIOR FILING DATE: 1998-06-10 |
| 15 | PRIOR APPLICATION NUMBER: 60/088742 |
| 16 | PRIOR FILING DATE: 1998-06-10 |
| 17 | PRIOR APPLICATION NUMBER: 60/088810 |
| 18 | PRIOR FILING DATE: 1998-06-10 |
| 19 | PRIOR APPLICATION NUMBER: 60/088824 |
| 20 | PRIOR FILING DATE: 1998-06-10 |
| 21 | PRIOR APPLICATION NUMBER: 60/088826 |
| 22 | PRIOR FILING DATE: 1998-06-10 |
| 23 | PRIOR APPLICATION NUMBER: 60/088858 |
| 24 | PRIOR FILING DATE: 1998-06-11 |
| 25 | PRIOR APPLICATION NUMBER: 60/088861 |
| 26 | PRIOR FILING DATE: 1998-06-11 |
| 27 | PRIOR APPLICATION NUMBER: 60/088876 |
| 28 | PRIOR FILING DATE: 1998-06-11 |
| 29 | PRIOR APPLICATION NUMBER: 60/089105 |
| 30 | PRIOR FILING DATE: 1998-06-12 |
| 31 | PRIOR APPLICATION NUMBER: 60/089440 |
| 32 | PRIOR FILING DATE: 1998-06-16 |
| 33 | PRIOR APPLICATION NUMBER: 60/089512 |
| 34 | PRIOR FILING DATE: 1998-06-16 |
| 35 | PRIOR APPLICATION NUMBER: 60/089514 |
| 36 | PRIOR FILING DATE: 1998-06-16 |
| 37 | PRIOR APPLICATION NUMBER: 60/089532 |
| 38 | PRIOR FILING DATE: 1998-06-17 |
| 39 | PRIOR APPLICATION NUMBER: 60/089538 |
| 40 | PRIOR FILING DATE: 1998-06-17 |
| 41 | PRIOR APPLICATION NUMBER: 60/089598 |
| 42 | PRIOR FILING DATE: 1998-06-17 |
| 43 | PRIOR APPLICATION NUMBER: 60/089599 |
| 44 | PRIOR FILING DATE: 1998-06-17 |
| 45 | PRIOR APPLICATION NUMBER: 60/089600 |
| 46 | PRIOR FILING DATE: 1998-06-17 |
| 47 | PRIOR APPLICATION NUMBER: 60/089653 |
| 48 | PRIOR FILING DATE: 1998-06-17 |
| 49 | PRIOR APPLICATION NUMBER: 60/089801 |
| 50 | PRIOR FILING DATE: 1998-06-18 |
| 51 | PRIOR APPLICATION NUMBER: 60/089907 |
| 52 | PRIOR FILING DATE: 1998-06-18 |
| 53 | PRIOR APPLICATION NUMBER: 60/089908 |
| 54 | PRIOR FILING DATE: 1998-06-18 |
| 55 | PRIOR APPLICATION NUMBER: 60/089947 |
| 56 | PRIOR FILING DATE: 1998-06-19 |
| 57 | PRIOR APPLICATION NUMBER: 60/089948 |
| 58 | PRIOR FILING DATE: 1998-06-19 |
| 59 | PRIOR APPLICATION NUMBER: 60/089952 |
| 60 | PRIOR FILING DATE: 1998-06-19 |
| 61 | PRIOR APPLICATION NUMBER: 60/090246 |
| 62 | PRIOR FILING DATE: 1998-06-22 |
| 63 | PRIOR APPLICATION NUMBER: 60/090252 |
| 64 | PRIOR FILING DATE: 1998-06-22 |
| 65 | PRIOR APPLICATION NUMBER: 60/090254 |
| 66 | PRIOR FILING DATE: 1998-06-22 |
| 67 | PRIOR APPLICATION NUMBER: 60/090349 |
| 68 | PRIOR FILING DATE: 1998-06-23 |
| 69 | PRIOR APPLICATION NUMBER: 60/090355 |
| 70 | PRIOR FILING DATE: 1998-06-23 |
| 71 | PRIOR APPLICATION NUMBER: 60/090429 |
| 72 | PRIOR FILING DATE: 1998-06-24 |
| 73 | PRIOR APPLICATION NUMBER: 60/090431 |

| | | |
|----|-------------------------------------|--|
| 1 | PRIOR FILING DATE: 1998-06-24 | |
| 2 | PRIOR APPLICATION NUMBER: 60/090435 | |
| 3 | PRIOR FILING DATE: 1998-06-24 | |
| 4 | PRIOR APPLICATION NUMBER: 60/090444 | |
| 5 | PRIOR FILING DATE: 1998-06-24 | |
| 6 | PRIOR APPLICATION NUMBER: 60/090445 | |
| 7 | PRIOR FILING DATE: 1998-06-24 | |
| 8 | PRIOR APPLICATION NUMBER: 60/090472 | |
| 9 | PRIOR FILING DATE: 1998-06-24 | |
| 10 | PRIOR APPLICATION NUMBER: 60/090535 | |
| 11 | PRIOR FILING DATE: 1998-06-24 | |
| 12 | PRIOR APPLICATION NUMBER: 60/090540 | |
| 13 | PRIOR FILING DATE: 1998-06-24 | |
| 14 | PRIOR APPLICATION NUMBER: 60/090542 | |
| 15 | PRIOR FILING DATE: 1998-06-24 | |
| 16 | PRIOR APPLICATION NUMBER: 60/090557 | |
| 17 | PRIOR FILING DATE: 1998-06-24 | |
| 18 | PRIOR APPLICATION NUMBER: 60/090676 | |
| 19 | PRIOR FILING DATE: 1998-06-25 | |
| 20 | PRIOR APPLICATION NUMBER: 60/090678 | |
| 21 | PRIOR FILING DATE: 1998-06-25 | |
| 22 | PRIOR APPLICATION NUMBER: 60/090690 | |
| 23 | PRIOR FILING DATE: 1998-06-25 | |
| 24 | PRIOR APPLICATION NUMBER: 60/090694 | |
| 25 | PRIOR FILING DATE: 1998-06-25 | |
| 26 | PRIOR APPLICATION NUMBER: 60/090695 | |
| 27 | PRIOR FILING DATE: 1998-06-25 | |
| 28 | PRIOR APPLICATION NUMBER: 60/090696 | |
| 29 | PRIOR FILING DATE: 1998-06-25 | |
| 30 | PRIOR APPLICATION NUMBER: 60/090862 | |
| 31 | PRIOR FILING DATE: 1998-06-26 | |
| 32 | PRIOR APPLICATION NUMBER: 60/090863 | |
| 33 | PRIOR FILING DATE: 1998-06-26 | |
| 34 | PRIOR APPLICATION NUMBER: 60/091360 | |
| 35 | PRIOR FILING DATE: 1998-07-01 | |
| 36 | PRIOR APPLICATION NUMBER: 60/091478 | |
| 37 | PRIOR FILING DATE: 1998-07-02 | |
| 38 | PRIOR APPLICATION NUMBER: 60/091544 | |
| 39 | PRIOR FILING DATE: 1998-07-01 | |
| 40 | PRIOR APPLICATION NUMBER: 60/091519 | |
| 41 | PRIOR FILING DATE: 1998-07-02 | |
| 42 | PRIOR APPLICATION NUMBER: 60/091626 | |
| 43 | PRIOR FILING DATE: 1998-07-02 | |
| 44 | PRIOR APPLICATION NUMBER: 60/091633 | |
| 45 | PRIOR FILING DATE: 1998-07-02 | |
| 46 | PRIOR APPLICATION NUMBER: 60/091978 | |
| 47 | PRIOR FILING DATE: 1998-07-07 | |
| 48 | PRIOR APPLICATION NUMBER: 60/091982 | |
| 49 | PRIOR FILING DATE: 1998-07-07 | |
| 50 | PRIOR APPLICATION NUMBER: 60/092182 | |
| 51 | PRIOR FILING DATE: 1998-07-09 | |

| | | | | |
|-----------------------|--------|--------------------|-------|-------------|
| Query Match | 33.5%; | Score 555.5; | DB 9; | Length 331; |
| Best Local Similarity | 36.2%; | Pred. No. 1.8e-46; | | |

| | | | |
|----|-----|--|-----|
| Qy | 6 | MRLMYICLLVGLGALCYFYSMYSLNP-----FKQSEVYKYKDGKGNFLKLPDTCR | 53 |
| Db | 15 | LRSLKWSLLLSLLSFFVMYMLSLPHNVIERVNMWKFYEYPIYRQDFHFTLREHSNCS | 74 |
| Qy | 54 | QTPPELVLLVTSSSHOKLAERMAIQTQWGRMKWKGQOLKTFPFLCITTTSSAAE---TKVED | 110 |
| Db | 75 | HQNPFLVILTSHPDSVKARQALRVTWGEKKSWMGVEVLTFLLGQEAEEKDKMLALSLE | 134 |
| Qy | 111 | QESORHGDIIOKDFLDVYYNLLTKTMGIGEWHRFCPQAFAVNMKTOSDMFINVDYTELL | 170 |
| Db | 135 | DEHLLYGDIIIRQDFLDYNNLLTKTIMAPRWVTEFCPNKYVMKTDVDFINTGLNVKYL | 194 |
| Qy | 171 | LKNRIRTFETGELKLNFPPIRQPFKSWFKSEYWDRIYPPPCSTGTGVFGSDVASQVY | 230 |
| Db | 195 | LNUNHSEKFTGPLINDISYRGFYOKTHLSYQEYPFKVPFPYPCSGLGIVMSRDLVPRIY | 254 |
| Qy | 231 | NVSKSVPIYKLEDFVGLCLERLINIRLEELHQSOTPFPGGLRFSVCLFRRIYACHFIKPR | 290 |

Db 255 EMGHVPKEFDYVGICLNLLKVNHIHPEDTNLFELYRIHLDVCQLRRVIAAHGFSSK 314
Qy 291 TLDDYQAL 299
Db 315 EIIIFWQVM 323

Search completed: April 14, 2003, 12:18:15
Job time : 17 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: April 4, 2003, 14:37:51 ; Search time 38 Seconds
(without alignments)
981.847 Million cell updates/sec

Title: US-09-914-152-1-copy_31_310

Perfect score: 1498

Sequence: 1 FKQSFYKKDGNFLKLPDT.....TLDDYQWALENSRGDCPPV 280

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

```

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 1498 | 100.0 | 310 | 21 | Human beta-1,3 gal |
| 2 | 1492 | 99.6 | 310 | 21 | Human beta3Gal-T5 |
| 3 | 1016 | 67.8 | 192 | 23 | Human 5174397/63-2 |
| 4 | 577 | 38.5 | 326 | 15 | Beta-1,3-galactosy |
| 5 | 569 | 38.0 | 422 | 20 | Polypeptide identifi |
| 6 | 548.5 | 36.6 | 319 | 20 | Human Dendriac. H |
| 7 | 548.5 | 36.6 | 331 | 21 | Membrane-bound pro |
| 8 | 548.5 | 36.6 | 331 | 22 | Human PRO polypt |
| 9 | 548.5 | 36.6 | 331 | 22 | Human PRO1074. HO |
| 10 | 548.5 | 36.6 | 331 | 22 | Human PRO1074 (UNQ |

| | | | | | | |
|----|-------|------|-----|----|-----------|--------------------|
| 11 | 548.5 | 36.6 | 331 | 23 | AAU80224 | Human beta1,3-acet |
| 12 | 543.5 | 36.3 | 331 | 22 | AAU93992 | Human protein sequ |
| 13 | 497.5 | 33.2 | 193 | 23 | AAE15927 | Human 4566767/84-3 |
| 14 | 497.5 | 33.2 | 193 | 23 | AAE15928 | Mouse 2745735/84-3 |
| 15 | 480 | 32.0 | 196 | 23 | AAE15929 | Human 4502339_JBC_ |
| 16 | 479 | 32.0 | 196 | 23 | AAE15930 | Mouse 2745737/157- |
| 17 | 443.5 | 29.6 | 195 | 23 | AAE15933 | Mouse 2745739/84-3 |
| 18 | 433.5 | 28.9 | 195 | 23 | AAE15932 | Human 4502343_JBC_ |
| 19 | 413 | 27.6 | 372 | 19 | AAW64558 | Human epidermoid c |
| 20 | 413 | 27.6 | 372 | 22 | AAU23167 | Human PRO polypt |
| 21 | 413 | 27.6 | 372 | 22 | AAAB88404 | Human membrane or |
| 22 | 413 | 27.6 | 372 | 22 | AAAB49749 | Human beta 1,3-N-a |
| 23 | 413 | 27.6 | 372 | 22 | AAAB49750 | Human beta 1,3-N-a |
| 24 | 413 | 27.6 | 372 | 23 | ABB09716 | Amino acid sequenc |
| 25 | 413 | 27.6 | 372 | 23 | AAU11272 | Human beta1,3-N-ac |
| 26 | 401.5 | 26.8 | 368 | 23 | AAU11273 | Murine beta1,3-N-a |
| 27 | 396 | 26.4 | 378 | 22 | AAE04752 | Human beta-1,3-gal |
| 28 | 396 | 26.4 | 378 | 23 | AAE21679 | Human galactosyltr |
| 29 | 396 | 26.4 | 378 | 23 | ABB06980 | Human beta 1,3-N-a |
| 30 | 396 | 26.4 | 378 | 23 | ABB07522 | Human drug metabol |
| 31 | 393 | 26.2 | 371 | 20 | AAW95171 | Protein exhibiting |
| 32 | 387.5 | 25.9 | 378 | 20 | AAU34171 | Human galactosyltr |
| 33 | 387.5 | 25.9 | 378 | 22 | AAW93908 | Human polypeptide, |
| 34 | 384 | 25.6 | 401 | 23 | AAE22148 | Human TRNFR-10 pro |
| 35 | 381.5 | 25.5 | 574 | 22 | ABB63976 | Drosophila melanog |
| 36 | 365.5 | 24.4 | 334 | 22 | ABG20251 | Novel human diagno |
| 37 | 360 | 24.0 | 373 | 22 | AAW93536 | Human protein sequ |
| 38 | 360 | 24.0 | 397 | 19 | AAW80212 | Human cardiac and |
| 39 | 360 | 24.0 | 397 | 21 | AAW84683 | Amino acid sequenc |
| 40 | 360 | 24.0 | 397 | 21 | AAW69698 | Human Brainiac pro |
| 41 | 360 | 24.0 | 397 | 22 | AAW66118 | Human LIG46 polypt |
| 42 | 360 | 24.0 | 413 | 22 | AAW49748 | Human beta 1,3-N-a |
| 43 | 359.5 | 24.0 | 402 | 21 | AAW24035 | Human PRO4397 prot |
| 44 | 359.5 | 24.0 | 402 | 23 | AAE22142 | Human TRNFR-4 prot |
| 45 | 359.5 | 24.0 | 402 | 23 | ABG34046 | Human pro peptide |

ALIGNMENTS

RESULT 1
AAB26749

ID AAB26749 standard; Protein; 310 AA.

AC AAB26749;

XX 15-JAN-2001 (first entry)

DE Human beta-1,3 galactose transferase protein sequence.

KW Beta-1,3 galactose transferase; treatment; diagnosis; cancer; human;
digestive system.

OS Homo sapiens.

PN WO200050608-A1.

PD 31-AUG-2000.

PF 24-FEB-2000; 2000WO-JP01070.

PR 25-FEB-1999; 95JP-0047571.

PA (KYOW) KYOWA HAKKO KOGYO KK.

PI Narimatsu H, Isshiki S, Togayachi A, Sasaki K;

DR WPI; 2000-549409/50.

PT N-PSDB; AAA93875.

PT Beta-1,3 galactose transferase and DNA encoding it, useful for
synthesis of type 1 sialyl Lewis, a carbohydrate for treatment of
digestive system cancer

| | |
|----|--|
| XX | Claim 2; Page 97-98; 123pp; Japanese. |
| XX | This invention relates to a polypeptide (I) with beta-1,3 galactose |
| XX | transferase activity, or variants of (I) comprising amino acid additions, |
| CC | deletions and/or substitutions. Included in the invention is DNA encoding, |
| CC | all or part of (I); expression vectors containing the DNA, host cells |
| CC | transformed by the vectors; a method for the preparation of the |
| CC | polypeptide by culture of the transformants or by expression in the milk |
| CC | of a transgenic mammal, and antibodies recognising (I). The Beta-1,3 |
| CC | galactose transferase protein transfers galactose by beta-1,3 bonding to |
| CC | N-acetylglucosamine present in a non-cyclic carbohydrate chain (such as |
| CC | GlcNAc beta1-3gal beta1-4Glc) to give Gal beta1-3GlcNAc. The protein and |
| CC | DNA encoding it are useful for the treatment and diagnosis of cancer of |
| CC | the digestive system. The present sequence represents the Beta-1,3 |
| CC | galactose transferase protein. |
| XX | |
| XX | Sequence 310 AA; |
| XX | |
| XX | Query Match 100.0%; Score 1498; DB 21; Length 310; |
| XX | Best Local Similarity 100.0%; Pred. No. 5.2e-165; |
| XX | Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |
| XX | |
| QY | 1 FKEQSVYKKGDFLKPDTDCRTPPELVLLVTSSSHKQLAERMAIRQTGKERVMVKGQ 60 |
| DB | 31 FKEQSVYKKGDFLKPDTDCRTPPELVLLVTSSSHKQLAERMAIRQTGKERVMVKGQ 90 |
| QY | 61 LKTFELGTTSSAAETKEVDQESORHGDIIOKDFLDVYYNLTLTMMGIEWVHRPCPOAA 120 |
| DB | 91 LKTFELGTTSSAAETKEVDQESORHGDIIOKDFLDVYYNLTLTMMGIEWVHRPCPOAA 150 |
| QY | 121 FVMKTDSDMFINVDTLTLLLLKKNRTTRFFTGFLKLNFFPIRQPSKWFVSKSEYPWDRY 180 |
| DB | 151 FVMKTDSDMFINVDTLTLLLLKKNRTTRFFTGFLKLNFFPIRQPSKWFVSKSEYPWDRY 210 |
| QY | 181 PPFCSGTGYVFGSDVASQVYVNSKVPYIKLEDVFGVGLCLERLNIRLEELHSQPTFFPGG 240 |
| DB | 211 PPFCSGTGYVFGSDVASQVYVNSKVPYIKLEDVFGVGLCLERLNIRLEELHSQPTFFPGG 270 |
| QY | 241 LRFVSVCLFRRIVACHFIKPRTLDDYQWALENSRGEDCCPV 280 |
| DB | 271 LRFVSVCLFRRIVACHFIKPRTLDDYQWALENSRGEDCCPV 310 |
| XX | |
| XX | RESULT 2 |
| XX | AA94641 |
| XX | ID AAY94641 standard; Protein; 310 AA. |
| XX | AA94641; |
| XX | |
| XX | 15-AUG-2000 (first entry) |
| XX | Human beta3Gal-T5 protein sequence. |
| XX | |
| XX | UDP-D-galactose; beta-N-acetylglucosamine beta1,3-galactosyltransferase; |
| XX | Beta3Gal-T5; human; chromosome 2/q22.3; galactosylation; glycoprotein; |
| XX | Beta1,3-galactosyl glycosylated saccharide production; glycopeptide. |
| XX | |
| XX | Homio sapiens. |
| XX | |
| XX | Key Location/Qualifiers |
| XX | Domain 13..26 |
| XX | /note= "Transmembrane domain" |
| XX | |
| XX | W0200029558-A1. |
| XX | |
| XX | 25-MAY-2000. |
| XX | |
| XX | 11-NOV-1999; 99WO-US26807. |
| XX | |
| XX | 13-NOV-1998; 98OK-0001483. |
| XX | |
| XX | {CLAU/ CLAUSEN H. |

| | |
|----------|---|
| XX | Clausen H, Amado M; |
| PI | WPI: 2000-399728/34. |
| DR | N-PSDB; AAA27959. |
| DR | |
| XX | |
| PT | Novel nucleic acid sequence encoding human |
| PT | UDP-galactose:beta-N-acetylglucosamine beta1,3-galactosyltransferase |
| PT | useful for obtaining beta 1,3-galactosyl glycosylated saccharides and |
| PT | glycopeptides or glycoproteins - |
| XX | |
| PS | Claim 23; Fig 1; 74pp; English. |
| XX | |
| CC | The present invention relates to a nucleic acid sequence encoding |
| CC | UDP-D-galactose:beta-N-acetylglucosamine beta1,3-galactosyltransferase |
| CC | (beta3gal-T5). Beta3 transferases add galactose to the hydroxy group at |
| CC | carbon 3 of 2-acetamido-2-deoxy-D-glucose (GlcNAc). The present sequence |
| CC | represents the human beta3gal-T5 protein sequence. The beta3gal-T5 gene |
| CC | is located on human chromosome 21q22.3. Beta3gal-T5 is a type II |
| CC | transmembrane glycoprotein. The invention also relates to the beta3gal-T5 |
| CC | protein sequence, a nucleic acid vector comprising the beta3gal-T5 |
| CC | nucleotide sequence, a host cell comprising the vector, and a method for |
| CC | the production of the beta3gal-T5 protein from the host cells. The |
| CC | methods of the invention can be used for recombinant production of |
| CC | beta3gal-T5 for use as a catalyst and for recombinant production of |
| CC | peptides or proteins with appropriate galactosylation. The beta3gal-T5 |
| CC | protein can be used to obtain beta1,3-galactosyl glycosylated |
| CC | saccharides, glycopeptides or glycoproteins. |
| XX | |
| SQ | Sequence 310 AA; |
| | Query Match 99.6%; Score 1492; DB 21; Length 310; |
| | Best Local Similarity 99.6%; Pred. No. 2.6e-164; |
| | Matches 279; Conservative 0; Mismatches 1; Indels 0; Gaps 0 |
| Qy | 1 FKESFVYKKDGNFLKLPDTCRTPPELVLLVTSSHKOLAERMAIRQTWGKERMYKGKQ 60 |
| Db | 31 FKESFVYKKDGNFLKLPDTCRTPPELVLLVTSSHKOLAERMAIRQTWGKERTVKGQ 90 |
| Qy | 61 LKTFPELLGTSSAETKEVDQESORRGDI IQKDLDVVYNLTLMKTMMGI EWVHRCPPQA 120 |
| Db | 91 LKTFPELLGTSSAETKEVDQESORRGDI IQKDLDVVYNLTLMKTMMGI EWVHRCPPQA 150 |
| Qy | 121 FVMKTDSDMF INV DYLT ELL KKNRTTRFFTGFLKLNEFFIRQPFSKMVF SKSEYPWDRY 180 |
| Db | 151 FVMKTDSDMF INV DYLT ELL KKNRTTRFFTGFLKLNEFFIRQPFSKMVF SKSEYPWDRY 210 |
| Qy | 181 PPFCSGTGYVSGDVASOVYVNSKVPIKLEDVFGVLCLERLNIRLELHSOPTFPFGG 240 |
| Db | 211 PPFCSGTGYVSGDVASOVYVNSKVPIKLEDVFGVLCLERLNIRLELHSOPTFPFGG 270 |
| Qy | 241 LRFSVCLFRRI VACHFIKPTRLIDYWQALNSRGEDCPPV 280 |
| Db | 271 LRFSVCLFRRI VACHFIKPTRLIDYWQALNSRGEDCPPV 310 |
| RESULT 3 | |
| AAE15931 | ID AAE15931 standard; Protein; 192 AA. |
| XX | |
| AC | AAE15931: |
| XX | |
| DT | 26-MAR-2002 (first entry) |
| XX | |
| DE | Human 5174397/63-284 protein, member of Fringe and Brainiac family. |
| KW | Fringe protein; Brainiac protein; glycosyltransferase; EGF; mannose; |
| KW | N-acetylglucosamine; fucose; epidermal growth factor; T cell leukaemia; |
| KW | Breast cancer; stroke; dementia; leukoencephalopathy; Alagille syndrome; |
| KW | notch protein; EGF-module containing protein; human. |
| OS | Homo sapiens. |
| XX | |

(PN WO2001187321-A2.
 XX 22-NOV-2001.
 XX 21-MAY-2001; 2001WO-IB01033.
 XX 19-MAY-2000; 2000GB-0012216.
 XX (EUMO-) EURO MOLECULAR BIOLOGY LAB.
 PA Cohen S, Brueckner K, Clausen H, Keck B;
 PI WPI; 2002-082940/11.
 XX
 XX Use of Fringe protein or Brainiac protein as glycosyltransferase and
 PT treatment of disease caused by epidermal growth factor-like module
 PT containing protein, such as T cell leukemia, breast cancer, stroke and
 PT dementia
 XX
 XX Claim 7; Fig 1; 52pp; English.
 PS
 XX The invention relates to the use of a Fringe protein or a Brainiac
 CC protein or a fragment or functional equivalent of a Fringe or a
 CC Brainiac protein, as glycosyltransferase. Fringe and Brainiac
 CC proteins have been found to possess glycosyltransferase activity in
 CC transferring sugar residues onto certain proteins of biological
 CC interest, so affecting the binding of effector molecules to these
 CC proteins. Particularly they are useful for transferring a
 CC N-acetylglucosamine moiety onto a fucose or a mannose substrate,
 CC whether free or attached to a lipid, carbohydrate or protein. They
 CC act as glycosyltransferase on notch protein (EGF-module containing
 CC protein) family. They are useful in treatment of a disease caused
 CC by epidermal growth factor (EGF)-like module containing protein,
 CC such as T cell leukaemia, breast cancer, stroke, dementia, cerebral
 CC autosomal dominant arteriopathy with subcortical infarcts,
 CC leucoencephalopathy and Alagille syndrome. Fringe and Brainiac
 CC proteins are also useful for screening ligands capable of modulating
 CC the activity of the protein. The present sequence is human protein
 CC which is a member of Fringe and Brainiac protein family.
 XX
 SQ Sequence 192 AA;

 Query Match 67.8%; Score 1016; DB 23; Length 192;
 Best Local Similarity 100.0%; Pred. No. 2.5e-109;
 Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 33 VTSSHKQLAERMAIROTGWGKRMVKGKQKLTFFLLGTTSSAAETKEVDQESORHGDIQK 92
 DB 1 VTSSHKQLAERMAIROTGWGKRMVKGKQKLTFFLLGTTSSAAETKEVDQESORHGDIQK 60

 QY 93 DFLDVYVNLTKTMGMIEWVHRFCPQAAFYMKTDSDMFINVYDYLTELLKKNRTTRFTG 152
 DB 61 DFLDVYVNLTKTMGMIEWVHRFCPQAAFYMKTDSDMFINVYDYLTELLKKNRTTRFTG 120

 QY 153 FLKLINEPIQPPSKVSEYKSEYVWDYPPFCSTGYVFGSDVASQVYVNSKSVYIKLE 212
 DB 121 FLKLINEPIQPPSKVSEYKSEYVWDYPPFCSTGYVFGSDVASQVYVNSKSVYIKLE 180

 QY 213 DVFVGLCLERLN 224
 DB 181 DVFVGLCLERLN 192

 RESULT 4
 AAR57433
 ID AAR57433 standard; Protein; 326 AA.
 XX
 AC AAR57433;
 XX
 DT 14-MAR-1995 (first entry)
 XX
 DE Beta-1,3-galactosyltransferase.
 XX

KW Beta-1,3-galactosyltransferase; enzyme; saccharide chain; pAMOPRWM1;
 KW KJM-1 cells.
 XX
 OS Homo sapiens.
 PN JP06181759-A.
 XX
 PD 05-JUL-1994.
 XX
 PF 16-DEC-1992; 92JP-0336436.
 XX
 PR 16-DEC-1992; 92JP-0336436.
 XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 XX WPI; 1994-251683/31.
 DR N-PSDB; AAR57433.
 XX
 PT Beta-galactosyl-transferase DNA and protein - useful for prodn. of
 PT saccharide chains
 XX
 PS Claim 1; Page 22-24; 47pp; Japanese.
 XX
 CC This sequence represents a beta-1,3-galactosyltransferase. This
 CC enzyme can be used to produce physiologically active saccharide
 CC chains and variants, and for improvement of saccharide chains bound
 CC to physiologically active proteins. The cDNA encoding this protein
 CC represents a fragment of the plasmid pAMOPRWM1 which was cloned in
 CC KJM-1 cells.
 XX
 SQ Sequence 326 AA;

 Query Match 38.5%; Score 577; DB 15; Length 326;
 Best Local Similarity 42.8%; Pred. No. 5.5e-58;
 Matches 116; Conservative 46; Mismatches 93; Indels 16; Gaps 4;

 QY 14 FLKLPDPTDCRTPPLVLLVTTSSHKQLAERMAIROTGWGKRMVKGKQKLTFFLLGTTSSA 73
 DB 65 FLKLPDPTDCRTPPLVLLVTTSSHKQLAERMAIROTGWGKRMVKGKQKLTFFLLGTTSSA 124

 QY 74 AETKEVDQESORHGDIQKDFLDVYVNLTKTMGMIEWVHRFCPQAAFYMKTDSDMFINV 133
 DB 125 VLNQWVEQESQIFHDIIVEDEFIDSYHNLTAKTLMGMRVAVATFCSKAKYVMTDSDIFVNM 184

 QY 134 DYLTELLKKNR--TTRFTGFLKLINEPIQPPSKVSEYKSEYVWDYPPFCSTGYV 191
 DB 185 DNLIYKLLKPKTPRRRYFTGYV-INGGPIRDVRSKWTMPRDLYPDSDNYPPFCSTGYIF 243

 QY 192 SGDVASQVYVNSKSVYIKLEDFVGLCLERLNIRLELHSDPTFFPGG-----LRFVSC 246
 DB 244 SADVAELIYKTSLHTRLLHLEDVYVGLCLERLGIH-----PTQNSGFNHWKWAYSILC 295

 QY 247 LFRIVACHFTKPRTLTDYWOALENSRGDC 277
 DB 296 RYRRVITVHQISPEEMHRIWMDSSKKHLRC 326

 RESULT 5
 AAY02366
 ID AAY02366 standard; Protein; 422 AA.
 XX
 AC AAY02366;
 XX
 DT 13-JUL-1999 (first entry)
 XX
 DE Polypeptide identified by the signal sequence trap method.
 XX
 KW Signal sequence trap method; SST method; immunisation; inhibition;
 KW infection; allergy; cancer; regulation; tissue formation; tissue repair;
 KW activin activity; inhibin activity; chemokine activity;
 KW cytokine activity; blood coagulation regulation; agonist; antagonist;
 KW metabolic disorder; hormonal disorder; immune disorder; wound;
 KW severe combined immunodeficiency; SCID; AIDS; thrombosis; cancer.

```

XX OS Homo sapiens.
XX PN W09918126-A1.
XX PD 15-APR-1999.
XX PF 06-OCT-1998; 98WO-JP04514.
XX PR 07-OCT-1997; 97JP-0274674.
XX PA (ONOH) ONO PHARM CO LTD.
XX PI Fukushima D, Shibayama S, Tada H;
XX DR WPI; 1999-277254/23.
XX DR N-PSDB; AAX35710, AAX35711.
XX PT Polypeptides identified by the signal sequence trap method from a
XX PT human cDNA library
XX PS Claim 1; Page 106-108; 28lpp; Japanese.
XX CC AAY02358-84 represent novel polypeptides which are identified from a
XX CC human placental cDNA library by the signal sequence trap (SST) method.
XX CC The polypeptides are encoded by the cDNA sequences in AAX35694-X35747.
XX CC The polypeptides have a broad range of physiological activity, including
XX CC immunisation against and inhibition of infections, allergies and cancer;
XX CC regulation of tissue formation and repair; activin/inhibin activity;
XX CC chemokine/cytokine activity; blood coagulation regulation; and
XX CC receptor/ligand agonist or antagonist activity. The polypeptides can
XX CC be used for prevention and treatment of disorders including infections
XX CC by bacteria, yeasts and viruses (including HIV) and protozoa; metabolic
XX CC and hormonal disorders; immune disorders (including severe combined
XX CC immunodeficiency (SCID) and AIDS; thrombosis; cancer; and traumatic or
XX CC surgical wounds.
XX SQ Sequence 422 AA;
    Query Match 38.0%; Score 569; DB 20; Length 422;
    Best Local Similarity 41.5%; Pred. No. 7e-57;
    Matches 108; Conservative 53; Mismatches 95; Indels 4; Gaps 2;
QY 22 CROTTPFLVLLTSSHKQLAERMAIROTWGKERMYKGLKTFLLGTT--SSAATKEV 79
Db 146 CQKSPFLLLTAEPGQTEARRAIQOTWGNESLAPGIQITRIFLGLSLKUNGYYLQRAI 205
QY 80 DQESQRHGDIIOKDFLDVYVYNTLTKTMGIEWHRECPQAAFYMKTDSMFINVDTLTEL 139
Db 206 LEESRQYHDIIOQEVLDVYVYNTLTKTMGMNVATVCPHIPVYMKTDSMFNVTEYLINK 265
QY 140 LLKKNTTR--FTGFLKLNPEPIROPFSKWFVSKSEYPDWDRYPPPCSGTGYVFGDVAS 197
Db 266 LLKPLDPPRRHFTGFLMRGYAPNRNKKSKWMPDLPYSERYPVFCSGTGYVFGDLAE 325
QY 198 QVYVNSKSVPIKLEDFVGLCLERLNIRLEELHSDPTFFPGGLRFSVCLFRRIVACHFI 257
Db 326 KIFKVLGIRRUHLEDDVYVGICLAKLRIDPVPPPEFVFNHWRVSYSCCKYSLHITSHQF 385
QY 258 KPTLLDYWQALENSRGDC 277
Db 386 QPSELIKYWNHLQONKHNAC 405
RESULT 6
AAY06461
ID AAY06461 standard; Protein; 319 AA.
XX AC AAY06461;
XX DT 27-SEP-1999 (first entry)
XX DE Human Dendriac.
```

```

XX KW Dendriac; Brainiac-2; human; Notch; immune disorder;
XX KW neurological disorder; diagnosis; therapy.
XX OS Homo sapiens.
XX FH Key
XX FH Peptide
FT 1..25 /note= "signal peptide"
FT 26..319 /note= "mature protein"
FT 60 /note= "N-glycosylated"
FT 142 /note= "N-glycosylated"
FT 186 /note= "N-glycosylated"
FT 200 /note= "N-glycosylated"
FT 314 /note= "N-glycosylated"
FT 82..90 /label= CD-I
FT /label= "conserved domain, shows high identity
FT to Brainiac-3 and Drosophila Brainiac"
FT Domain
FT 100..109 /label= CD-II
FT /note= "conserved domain, shows high identity
FT to Brainiac-3 and Drosophila Brainiac"
FT Domain
FT 120..153 /label= CD-III
FT /note= "conserved domain, shows high identity
FT to Brainiac-3 and Drosophila Brainiac"
FT Domain
FT 168..182 /label= CD-IV
FT /note= "conserved domain, shows high identity
FT to Brainiac-3 and Drosophila Brainiac"
FT Domain
FT 199..204 /label= CD-V
FT /note= "conserved domain, shows high identity
FT to Brainiac-3 and Drosophila Brainiac"
FT Domain
FT 217..225 /label= CD-VI
FT /note= "conserved domain, shows high identity
FT to Brainiac-3 and Drosophila Brainiac"
FT Domain
FT 234..242 /label= CD-VII
FT /note= "conserved domain, shows high identity
FT to Brainiac-3 and Drosophila Brainiac"
FT Domain
FT 250..261 /label= CD-VIII
FT /note= "conserved domain, shows high identity
FT to Brainiac-3 and Drosophila Brainiac"
FT Domain
FT 274..282 /label= CD-IX
FT /note= "conserved domain, shows high identity
FT to Brainiac-3 and Drosophila Brainiac"
FT Domain
FT 285..291 /label= CD-X
FT /note= "conserved domain, shows high identity
FT to Brainiac-3 and Drosophila Brainiac"
FT Domain
FT 297..308 /label= CD-XI
FT /note= "conserved domain, shows high identity
FT to Brainiac-3 and Drosophila Brainiac"
FT Peptide
FT 55..63 /note= "epitope-bearing peptide"
FT Peptide
FT 75..84 /note= "epitope-bearing peptide"
FT Peptide
FT 89..95 /note= "epitope-bearing peptide"
FT Peptide
FT 233..238 /note= "epitope-bearing peptide"
```

[illegible]

| | | | |
|-----------------------|----------|---|-----|
| Qy | 118 | QAFAVVKMTSDMPINVDYITELLKKNRTTRFTTGFGLKLNFFPIRQPPSKWFSKSEYWP | 177 |
| Db | 172 | NAXYVVKMTDTPVINTGNLVKYLNLNHNSEKFFTGYPIDINYSYRGFYQKTH1SYQYPF | 231 |
| Qy | 178 | DRYPFCSGTGYVFGSDVASQVYNNYSKSVPIYIKLEDVFGVGLCLERLNRLEELHSQPTFF | 237 |
| Db | 232 | KVFPFCSGLGYIMSRDLVPRIYEMGHVKPIKFDVTVGICLNLLKVNIIHPEDTNLFF | 291 |
| Qy | 238 | PGGLRSVFCFRRIRIVACHFIKPRTLDDYQWAL | 269 |
| Db | 292 | LYRIHLDCQLRRVIAAHGFSKKEITTFQVM | 323 |
| RESULT 11 | | | |
| AAU80224 | | | |
| ID | AAU80224 | standard; Protein; 331 AA. | |
| XX | AC | AAU80224; | |
| XX | XX | | |
| DT | XX | 30-JUL-2002 (first entry) | |
| DE | XX | Human beta1,3-acetylglactosamine transferase. | |
| XX | XX | | |
| KW | XX | Human; enzyme; beta1,3-acetylglactosamine transferase; | |
| KW | XX | Gb4 sugar chain. | |
| XX | XX | | |
| OS | XX | Homo sapiens. | |
| XX | XX | | |
| PN | XX | JP2002085069-A. | |
| XX | XX | | |
| PD | XX | 26-MAR-2002. | |
| XX | XX | | |
| PF | XX | 08-SEP-2000; 2000JP-0273835. | |
| XX | XX | | |
| PR | XX | 08-SEP-2000; 2000JP-0273835. | |
| XX | XX | | |
| PA | XX | (SEKG) SEIKAGAKU KOGYO CO LTD. | |
| PA | XX | (FURU/) FURUKAWA K. | |
| XX | XX | | |
| DR | XX | WPI: 2002-378274/41. | |
| XX | XX | N-PSDB; ABK51201. | |
| XX | XX | | |
| PT | XX | Preparation of beta-1,3-acetylglactosamine transferase - | |
| XX | XX | | |
| PS | XX | Claim 3; Page 13-14; 15pp; Japanese. | |
| XX | XX | | |
| CC | CC | The invention relates to the preparation of beta-1,3-acetylglactosamine | |
| CC | CC | (a Gb4 sugar chain molecule) comprising introducing a DNA appearing as | |
| CC | CC | ABK51201 (S1) or a DNA hybridising with S1 or a base sequence | |
| CC | CC | complementary to the base sequence or part of these base sequences under | |
| CC | CC | a stringent condition into a cell and growing the cell to express beta-1, | |
| CC | CC | 3-acetylglactosamine transferase and collecting it. Also included is the | |
| CC | CC | preparation of Gb4 sugar chain comprising contacting a polypeptide | |
| CC | CC | comprising the beta-1,3-acetylglactosamine transferase protein appearing | |
| CC | CC | as AAU80224 (A) or a polypeptide consisting of an amino acid sequence in | |
| CC | CC | which at least one amino acid is replaced, deleted, inserted or | |
| CC | CC | transferred in the amino acid sequence (A) and having enzymatic activity | |
| CC | CC | transferring N-acetylglactosamine residue from an N-acetylglactosamine | |
| CC | CC | donor to the C3 site of the galactose residue in Gb3 sugar chain which is | |
| CC | CC | the receptor with an N-acetylglactosamine donor and Gb3 sugar chain. | |
| CC | CC | The present sequence represents beta-1,3-acetylglactosamine transferase. | |
| XX | XX | | |
| SQ | Sequence | 331 AA: | |
| Query Match | | 36.6%; Score 548.5; DB : Length 331; | |
| Best Local Similarity | | 38.6%; Pred. No. 1.2e-54; | |
| Matches 105; | | Conservative 53; Mismatches 111; Indels 3; Gaps 1 | |
| Qy | 1 | FREQSFVYKKDNFLKPLDTCROTPPFLVLLVSSHQKLAERNAIFQTWCKRMYKQ | 60 |
| Db | 52 | FYEYEPYRQDPHFILREHNSCHONPELVILVTSFSDVKARQAIRVTWGKKSWMGYE | 111 |
| Qy | 61 | LKTFLLGTTSSAAE---TKEYDQESQRHGDIIOKDFLDVYVNIUTLKTMMGIEWVHRFCP | 117 |

PT containing protein, such as T cell leukemia, breast cancer, stroke and
XX dementia
PS Claim 7; Fig 1; 52pp; English.
XX
CC The invention relates to the use of a Fringe protein or a Brainiac
CC protein or a fragment or functional equivalent of a Fringe or a
CC Brainiac protein, as glycosyltransferase. Fringe and Brainiac
CC proteins have been found to possess glycosyltransferase activity in
CC transferring sugar residues onto certain proteins of biological
CC interest, so affecting the binding of effector molecules to these
CC proteins. Particularly they are useful for transferring a
CC N-acetylglucosamine moiety onto a fucose or a mannose substrate,
CC whether free or attached to a lipid, carbohydrate or protein. They
CC act as glycosyltransferase on notch protein (EGF-module containing
CC protein) family. They are useful in treatment of a disease caused
CC by epidermal growth factor (EGF)-like module containing protein,
CC such as T cell leukemia, breast cancer, stroke, dementia, cerebral
CC autosomal dominant arteriopathy with subcortical infarcts,
CC leucoencephalopathy and Alagille syndrome. Fringe and Brainiac
CC proteins are also useful for screening ligands capable of modulating
CC the activity of the protein. The present sequence is mouse protein
CC which is a member of Fringe and Brainiac protein family.
XX
SQ Sequence 193 AA;
Query Match 33.2%; Score 497.5; DB 23; Length 193;
Best Local Similarity 49.7%; Pred. No. 4.3e-49;
Matches 96; Conservative 34; Mismatches 60; Indels 3; Gaps 2;
QY 33 VTSSHKQLAERMAIRQTGWKRWYKQKLTFFLLGTTSSAAETKEVDQESQRHGDIIOK 92
DB 1 ISTTHKEFDARQAIRETWGDENNEFKIKIATFLLLGNADPNVLMQWVEQESQIFHDIIVE 60
QY 93 DFLDYVNLTKTMGIEWVHRFCPPQAAFMVTKTDSDFINVDYLTLLKKNR--TTRFF 150
DB 61 DFIDSYNHLTKLTMGMNRVATFCGKAKYVMKTDSDIFVNMNDLIYRLKKPSTKPRRYF 120
QY 151 TGFLLKNEFFPIQPFKSWFVSKSEYPMDRYPFCSGTGYVFGDVASQVYVSKSPYIK 210
DB 121 TGYV-INGPIRDVRSKWMYPRDLYPDSDNYPFCSGTGYIFSADVAELIYKTSLHTRLH 179
QY 211 LEDVFVGLCLERL 223
DB 180 LEDVYVGLCLRLK 192
RESULT 15
AAE15929
ID AAE15929 standard; Protein; 196 AA.
XX
AC AAE15929;
XX
XX
DT 26-MAR-2002 (first entry)
XX
DE Human 4502339_JBC_1/157-364 protein member of Fringe and Brainiac family.
XX
KW Fringe protein; Brainiac protein; glycosyltransferase; EGF; mannose;
KW N-acetylglucosamine; fucose; epidermal growth factor; T cell leukemia;
KW breast cancer; stroke; dementia; leucoencephalopathy; Alagille syndrome;
KW notch protein; EGF-module containing protein; human.
XX
OS Homo sapiens.
XX
XX WO200187321-A2.
PN
XX
PD 22-NOV-2001.
XX
XX 21-MAY-2001; 2001WO-IB01033.
PF
XX 19-MAY-2000; 2000GB-0012216.
PR
XX
XX (EUMO-) EURO MOLECULAR BIOLOGY LAB.

XX
PI Cohen S, Brueckner K, Clausen H, Keck B;
XX WPI; 2002-082940/11.
XX
PT Use of Fringe protein or Brainiac protein as glycosyltransferase and
PT treatment of disease caused by epidermal growth factor-like module
PT containing protein, such as T cell leukemia, breast cancer, stroke and
XX dementia
PS Claim 7; Fig 1; 52pp; English.
XX
CC The invention relates to the use of a Fringe protein or a Brainiac
CC protein or a fragment or functional equivalent of a Fringe or a
CC Brainiac protein, as glycosyltransferase. Fringe and Brainiac
CC proteins have been found to possess glycosyltransferase activity in
CC transferring sugar residues onto certain proteins of biological
CC interest, so affecting the binding of effector molecules to these
CC proteins. Particularly they are useful for transferring a
CC N-acetylglucosamine moiety onto a fucose or a mannose substrate,
CC whether free or attached to a lipid, carbohydrate or protein. They
CC act as glycosyltransferase on notch protein (EGF-module containing
CC protein) family. They are useful in treatment of a disease caused
CC by epidermal growth factor (EGF)-like module containing protein,
CC such as T cell leukemia, breast cancer, stroke, dementia, cerebral
CC autosomal dominant arteriopathy with subcortical infarcts,
CC leucoencephalopathy and Alagille syndrome. Fringe and Brainiac
CC proteins are also useful for screening ligands capable of modulating
CC the activity of the protein. The present sequence is human protein
CC which is a member of Fringe and Brainiac protein family.
XX
SQ Sequence 196 AA;
Query Match 32.0%; Score 480; DB 23; Length 196;
Best Local Similarity 48.1%; Pred. No. 4.8e-47;
Matches 91; Conservative 36; Mismatches 58; Indels 4; Gaps 2;
QY 39 QLAERMAIRQTGWKRWYKQKLTFFLLGTT--SSAAETKEVDQESQRHGDIIOKDFLD 96
DB 7 QIARRAIRQTGWGNEGLAPGQITRIFLLGLSLKINGYLQRAILEESROYHDIIOQEYLD 66
QY 97 VYVNLTKTMGIEWVHRFCPPQAAFMVTKTDSDFINVDYLTLLKKNRTR--FFTGFL 154
DB 67 TYVNLTKLTMGMNRVATFCGKAKYVMKTDSDMFVNTYELNRLKKPDLPPRHVFTGYL 126
QY 155 KLNFFPIQPFKSWFVSKSEYPMDRYPFCSGTGYVFGDVASQVYVSKSPYIKLEDV 214
DB 127 MRGYAPNRNKDSKWMYPPDLYPSEYVFCSGTGYVFGDLAEKIFKVSGLIRRLHLEDV 186
QY 215 FVGLCLERL 223
DB 187 YVGICLAKL 195
Search completed: April 4, 2003, 14:44:08
Job time : 40 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2003, 14:40:30 ; Search time 11 seconds
(without alignments)
1055.761 Million cell updates/sec

Title: US-09-914-152-1_COPY_31_310

Perfect score: 1498

Sequence: 1 FKQSFVYKKDGNFLKLPDT.....TLDDYQWALENSRGDCPPV 280

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SWISSProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|---------------|--------------------|
| 1 | 1498 | 100.0 | 310 | 1 B3G5_HUMAN | O9Y2C3 h beta-1,3- |
| 2 | 1433 | 95.7 | 300 | 1 B3G5_GORGO | O9n293 g beta-1,3- |
| 3 | 1422 | 94.9 | 301 | 1 B3G5_PANPA | O9n294 p beta-1,3- |
| 4 | 1407 | 93.9 | 297 | 1 B3G5_PANTR | O9n295 p beta-1,3- |
| 5 | 1133.5 | 75.7 | 308 | 1 B3G5_MOUSE | O9j167 m beta-1,3- |
| 6 | 413 | 27.6 | 372 | 1 B3G8_HUMAN | O9Y289 h probable |
| 7 | 360 | 24.0 | 398 | 1 B3G7_HUMAN | O9nY97 h beta-1,3- |
| 8 | 105.5 | 7.0 | 1049 | 1 VP35_YEAST | O07468 saccharomyc |
| 9 | 93.5 | 6.2 | 336 | 1 MTHB_HABPH | P50193 haemophilus |
| 10 | 90.5 | 6.0 | 1314 | 1 SS22_YEAST | P25390 saccharomyc |
| 11 | 89.5 | 6.0 | 786 | 1 TLR1_HUMAN | O15399 homo sapien |
| 12 | 88.5 | 5.9 | 579 | 1 YHV0_YEAST | P38848 saccharomyc |
| 13 | 86.5 | 5.8 | 516 | 1 YVNF_AZOCH | P24423 azotobacter |
| 14 | 85 | 5.7 | 1073 | 1 HSER_PIG | P55204 sus scrofa |
| 15 | 84.5 | 5.6 | 711 | 1 ETF2_MXXVL | O9q8K4 myxoma viru |
| 16 | 84 | 5.6 | 306 | 1 YFV9_SCHPO | O9p7q6 schizosacch |
| 17 | 83 | 5.5 | 275 | 1 ISPH_THEMEA | O9x1f7 thermotoga |
| 18 | 82.5 | 5.5 | 525 | 1 DIMH_CAEEL | O17397 caenorhabdi |
| 19 | 82.5 | 5.5 | 713 | 1 PMIP_HUMAN | O99797 homo sapien |
| 20 | 82 | 5.5 | 576 | 1 HXT4_YEAST | P32467 saccharomyc |
| 21 | 81.5 | 5.4 | 359 | 1 UNG_YEAST | P12887 saccharomyc |
| 22 | 81 | 5.4 | 397 | 1 O22A_DROME | P81909 drosophila |
| 23 | 81 | 5.4 | 1042 | 1 SYI_BORBU | O51773 borrelia bu |
| 24 | 80.5 | 5.4 | 3386 | 1 POLG_DEN4 | P09866 d genome po |
| 25 | 79.5 | 5.3 | 612 | 1 BIR3_MOUSE | Q62210 mus musculu |
| 26 | 79.5 | 5.3 | 691 | 1 Y104_YEAST | P40460 saccharomyc |
| 27 | 78.5 | 5.2 | 757 | 1 METE_PASMU | P57843 pasteurella |
| 28 | 78.5 | 5.2 | 1002 | 1 YA37_ANASP | P58612 anabaena sp |
| 29 | 78.5 | 5.2 | 4377 | 1 ANK3_HUMAN | O12955 homo sapien |
| 30 | 78 | 5.2 | 469 | 1 LEU2_BUCAP | Q85065 buchnera ap |
| 31 | 78 | 5.2 | 519 | 1 IG31_METJA | O59025 methanococc |
| 32 | 78 | 5.2 | 618 | 1 BIR3_HUMAN | Q13490 homo sapien |
| 33 | 77.5 | 5.2 | 271 | 1 Y398_METJA | O57841 methanococc |

| | | | | | |
|----|------|-----|------|--------------|--------------------|
| 34 | 77.5 | 5.2 | 710 | 1 PMIP_RAT | Q01992 rattus norv |
| 35 | 77 | 5.1 | 246 | 1 Y402_BUCAL | P57482 buchnera ap |
| 36 | 77 | 5.1 | 741 | 1 GNT5_HUMAN | Q09328 homo sapien |
| 37 | 76.5 | 5.1 | 397 | 1 OM1C_CHLTR | P08780 chlamydia t |
| 38 | 76.5 | 5.1 | 711 | 1 ETF2_SFVKA | Q9q8y2 shope fibro |
| 39 | 76.5 | 5.1 | 977 | 1 KIT_BOVIN | P43481 bos taurus |
| 40 | 76 | 5.1 | 374 | 1 GLY1_CANAL | O13427 candida alb |
| 41 | 76 | 5.1 | 2077 | 1 TSGU_HSVGU | P52340 human herpe |
| 42 | 75 | 5.0 | 333 | 1 Y642_METJA | Q58059 methanococc |
| 43 | 75 | 5.0 | 454 | 1 Y662_HAEIN | P45217 haemophilus |
| 44 | 75 | 5.0 | 467 | 1 XANB_XANCP | P29956 x xanthan b |
| 45 | 75 | 5.0 | 494 | 1 DCE2_ARATH | Q42472 arabidopsis |

ALIGNMENTS

RESULT 1
B3G5_HUMAN STANDARD; PRT; 310 AA.
AC O9Y2C3; Q9NY96; Q9P1X7; Q9P1X6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-1,3-galactosyltransferase 5 (EC 2.4.1.-) (Beta-1,3-GalTase 5)
DE (Beta3Gal-T5) (b3Gal-T5) (UDP-galactose:beta-N-acetylglucosamine beta-1,3-galactosyltransferase 5) (UDP-Gal:beta-GlcNAc beta-1,3-galactosyltransferase 5) (Beta-3-Gx-T5).
GN B3GALT5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Colorectal adenocarcinoma;
RX MEDLINE=99230269; PubMed=10212226;
RA Ishiki S., Togayachi A., Kudo T., Nishihara S., Watanabe M., Kubota T., Kitajima M., Shiraishi N., Sasaki K., Andoh T., Narimatsu H.;
RA "Cloning, expression, and characterization of a novel UDP-galactose:beta-N-acetylglucosamine beta1,3-galactosyltransferase (beta3Gal-T5) responsible for synthesis of type1 chain in colorectal and pancreatic epithelia and tumor cells derived therefrom.;"
RL J. Biol. Chem. 274:12499-12507(1999).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=9937698; PubMed=10406968;
RA Zhou D., Berger E.G., Hennet T.;
RT "Molecular cloning of a human UDP-galactose:GlcNAc:beta1,3GalNAc beta1,3 galactosyltransferase gene encoding an O-linked core3-elongation enzyme.;"
RL Eur. J. Biochem. 263:571-576(1999).
[3]
RP SEQUENCE FROM N.A.
RX Anado M., Carneiro F., Clausen H.;
RT "Cloning and expression of two beta-1,3-galactosyltransferases: beta3Gal-T5 and beta3Gal-T6.;"
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D., Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W., Rosenthal A., Kudo H., Shibuya K., Kawasaki K., Asakawa S., Shintani A., Sasaki T., Nagamine K., Mitsuyma S., Antonarakis S.E., Minoshima S., Shimizu N., Nordstiek G., Hornischer K., Brandt P., Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Blocker H., Ramser J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E., Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,

RA Lehrach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
RN [5]
RP SEQUENCE OF 1-298 FROM N.A.
RA Liu Y., Saitou N.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RX MEDLINE=20047730; PubMed=10580128;
RA Amado M., Almeida R., Schwientek T., Clausen H.;
RT "Identification and characterization of large galactosyltransferase
gene families: galactosyltransferases for all functions.";
RL Biochim. Biophys. Acta 1473:35-53(1999).
CC -!- FUNCTION: CATALYZES THE TRANSFER OF GAL TO GLCNAC-BASED ACCEPTORS
CC WITH A PREFERENCE FOR THE CORE3 O-LINKED GLYCAN
CC GLCNAC(BETAL 3)GALNAC STRUCTURE. CAN USE GLYCOPOLIPID LC3CER AS AN
CC EFFICIENT ACCEPTOR
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (Potential).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN STOMACH, JEJUNUM, COLON,
CC PANCREAS, SMALL INTESTINE, TESTIS AND GASTROINTESTINAL AND
CC PANCREATIC CANCER CELL LINES. HARDLY DETECTED IN LUNG, LIVER,
CC ADRENAL GLAND AND PERIPHERAL BLOOD LEUKOCYTES.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 31
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB020337; BAA77664.1; -
CC EMBL: AF145784; AAF07880.1; -
CC EMBL: AJ006078; CAB91547.1; -
CC EMBL: AL163280; CAB90446.1; -
CC EMBL: AB041412; BAA94497.1; -
CC EMBL: AB041413; BAA94498.1; -
CC EMBL: AB041414; BAA94499.1; -
CC MIM: 604066; -
CC InterPro: IPR002659; Galactosyl_T.
CC DR Pfam: PF01762; Galactosyl_T; 1
CC Transferrase; Glycosyltransferase; Glycoprotein; Transmembrane;
CC Signal-anchor; Golgi stack; Multigene family.
CC DOMAIN 1 7 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 8 28 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC (POTENTIAL).
CC FT DOMAIN 29 310 LUMENAL, CATALYTIC (POTENTIAL).
CC FT CARBOHYD 130 130 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 174 174 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 231 231 F -> Y (IN REF. 3).
CC FT CONFLICT 26 26 Y -> N (IN REF. 3).
CC FT CONFLICT 42 42 G -> V (IN REF. 5; BAA94497).
CC FT CONFLICT 85 85 M -> T (IN REF. 3 AND 5; BAA94497).
CC FT CONFLICT 114 114 Q -> R (IN REF. 5; BAA94497).
CC FT CONFLICT 134 134 K -> E (IN REF. 5; BAA94497).
CC FT CONFLICT 234 234 L -> R (IN REF. 5; BAA94497).
CC FT CONFLICT 277 277 LDYQW -> WTTGR (IN REF. 5; BAA94498).
CC FT CONFLICT 293 297
CC SEQUENCE 310 AA; 36189 MW; 4DD7A19E3E648AA9 CRC64;
CC
CC Query Match 100.0%; Score 1498; DB 1; Length 310;
CC Best Local Similarity 100.0%; Pred. No. 2.6e-135;
CC Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC 1 FKEQSFVYKKGDNFLKLPDTCROTPTPFLVLLVTSKHKQLAERMAIRQTGKERMVKGKQ 60
CC |
CC 31 FKEQSFVYKKGDNFLKLPDTCROTPTPFLVLLVTSKHKQLAERMAIRQTGKERMVKGKQ 90
CC |
CC 61 LKTFLLGTTSSAAETKEVDQESQRHGDIIQKDFLDVYVNLTKTMGIEWVHRFCPOAA 120
CC |

Db 91 LKTFLLGTTSSAAETKEVDQESQRHGDIIQKDFLDVYVNLTKTMGIEWVHRFCPOAA 150
QY 121 FVMTKTDSDMFINVYDLTELLKKNKRTTFTGFLKLINEPPIROPFSKFWFSKSEYPDWRY 180
|
Db 151 FVMTKTDSDMFINVYDLTELLKKNKRTTFTGFLKLINEPPIROPFSKFWFSKSEYPDWRY 210
QY 181 PPFCSGTGYVFGSDVASQVYVSKSVYIKLEDVFGVGLCLERLNIRLEELHSOPTFFPGG 240
|
Db 211 PPFCSGTGYVFGSDVASQVYVSKSVYIKLEDVFGVGLCLERLNIRLEELHSOPTFFPGG 270
QY 241 LRFVCLFPRRIIVACHIFIKPRTLLDYQWALENSKRGDCPPV 280
|
Db 271 LRFVCLFPRRIIVACHIFIKPRTLLDYQWALENSKRGDCPPV 310
RESULT 2
B3G5 GORGO STANDARD; PRT: 300 AA.
AC Q9N293;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-1,3-galactosyltransferase 5 (EC 2.4.1.-) (Beta-1,3-GalTase 5)
DE (Beta3Gal-T5) (b3Gal-T5) (UDP-galactose:beta-N-acetylgalactosamine beta-
DE 1,3-galactosyltransferase 5) (UDP-Gal:beta-GlcNAc beta-1,3-
DE galactosyltransferase 5) (Beta-3-Gal-T5) (Fragment).
GN B3GALT5.
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9595;
ID LIU_Y_Saitou N.;
RP SEQUENCE FROM N.A.
RA Liu Y., Saitou N.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CATALYZES THE TRANSFER OF GAL TO GLCNAC-BASED ACCEPTORS
CC WITH A PREFERENCE FOR THE CORE3 O-LINKED GLYCAN
CC GLCNAC(BETAL 3)GALNAC STRUCTURE. CAN USE GLYCOPOLIPID LC3CER AS AN
CC EFFICIENT ACCEPTOR (BY SIMILARITY).
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (Potential).
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 31.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB041416; BAA94501.1; -
CC InterPro: IPR002659; Galactosyl_T.
CC Pfam: PF01762; Galactosyl_T; 1
CC Transferrase; Glycosyltransferase; Glycoprotein; Transmembrane;
CC Signal-anchor; Golgi stack; Multigene family.
CC DOMAIN 1 7 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 8 28 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC (POTENTIAL).
CC FT DOMAIN 29 >300 LUMENAL, CATALYTIC (POTENTIAL).
CC FT CARBOHYD 130 130 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 174 174 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 231 231 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT NON_TER 300 300
CC SEQUENCE 300 AA; 35104 MW; 1AA11692ED9F06FA CRC64;
CC
CC Query Match 95.7%; Score 1433; DB 1; Length 300;
CC Best Local Similarity 99.6%; Pred. No. 3.8e-129;
CC Matches 269; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CC
CC 1 FKEQSFVYKKGDNFLKLPDTCROTPTPFLVLLVTSKHKQLAERMAIRQTGKERMVKGKQ 60
CC |
CC 31 FKEQSFVYKKGDNFLKLPDTCROTPTPFLVLLVTSKHKQLAERMAIRQTGKERTVAGKQ 90
CC |

| | | | | | | | |
|----|-----|--------------------|------------|------------|-------------|--------------------|-----|
| Qy | 1 | FKESFYVKDGNFLKLPD | TDRCQT | PPFLVLLVTS | SHKQLAERMA | IRQTWGKERMVYKGO | 60 |
| Db | 31 | FKESFYVKDGNFLKLPD | TDRCQT | PPFLVLLVTS | SHKQLAERMA | IRQTWGKERTYKGO | 90 |
| Qy | 61 | LKTEFLGTTSSAAETKEV | DQESORHGDI | IKQFDLVYV | NLTKTMGIEWH | RFCPQAA | 120 |
| Db | 91 | LKTEFLGTTSSAAETKEV | DQESORHGDI | IKQFDLGYV | NLTKTMGIEWH | RFCPQAA | 150 |
| Qy | 121 | FVMKTDSDMF | INV DYLT | ELLKKNRTR | PTFTGFLK | NEPPIRQPSKFWFSKSEY | 180 |
| Db | 151 | FVMKTDSDMF | INV DYLT | ELLKKNRTR | PTFTGFLK | NEPPIRQPSKFWFSKSEY | 210 |
| Qy | 181 | PPFCSGTGYV | FGDVASQV | YNVSKSVPY | IKLEDVFG | VLCLERLNIRLEELH | 240 |
| Db | 211 | PPFCSGTGYV | FGDVASQV | YNVSESVPY | IKLEDVFG | VLCLERLNIRLEELH | 270 |

```

Db      271  LRFVCFRRIVACHFIKPTLLDYWOALEN 301
|||||
RESULT 4
B3G5_PANTR
ID      B3G5_PANTR      STANDARD;      PRT;      297  AA.
AC      Q9N295;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Beta-1,3-galactosyltransferase 5 (EC 2.4.1.1-) (Beta-1,3-GalTase 5)
DE      (Beta3Gal-T5) (b3gal-T5) (UDP-galactose:beta-N-acetylgalucosamine beta
DE      1,3-galactosyltransferase 5) (UDP-Gal:beta-GlcNAc beta-1,3-
DE      galactosyltransferase 5) (Beta-3-Gx-T5) (Fragment).
GN      B3GALT5.
NC      Pan troglodytes (Chimpanzee).
OS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
NC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX      NCBI_TaxID=9598;
RN      [1]
RA      SEQUENCE FROM N.A.
RP      Liu Y., Saitou N.;
RL      Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC      -1- FUNCTION: CATALYZES THE TRANSFER OF GAL TO GLCNAC-BASED ACCEPTORS
CC      WITH A PREFERENCE FOR THE CORE3 O-LINKED GLYCAN
CC      GLCNAC(BETA1,3)GALNAc STRUCTURE. CAN USE GLYCOLIPID LC3CER AS AN
CC      EFFICIENT ACCEPTOR (BY SIMILARITY).
CC      -1- PATHWAY: Glycosylation.
CC      -1- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (Potential)
CC      -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 31.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collabora
CC      between the Swiss Institute of Bioinformatics and the EMBL outstat
CC      European Bioinformatics Institute. There are no restrictions on the
CC      use of this file.
CC      -----

```

| | | | |
|----|--|--|--|
| CC | | EMBL; AB041414; BAA94499.1; - | |
| DR | | InterPro: IPR002659; Galactosyl_T. | |
| DR | | Pfam: PF01762; Galactosyl_T; 1. | |
| KW | | Transferase; Glycosyltransferase; Glycoprotein; Transmembrane; | |
| FT | | Signal-anchor; Golgi stack; Multigene family. | |
| KF | | DOMAIN 1 7 CYTOPLASMIC (POTENTIAL). | |
| FT | | TRANSMEM 8 28 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) | |
| FT | | (POTENTIAL). | |
| FT | | DOMAIN 29 >29 LUMENAL, CATALYTIC (POTENTIAL). | |
| FT | | CARBOHYD 130 130 N-LINKED (GLCNAC. .) (POTENTIAL). | |
| FT | | CARBOHYD 174 174 N-LINKED (GLCNAC. .) (POTENTIAL). | |

Query Match 93.9%; Score 1407; DB 1; Length 297;

[illegible]

| | | | | | | |
|---|--|---|-----------|---|---|---|
| KW | Transferase; Glycosyltransferase; Glycoprotein; Transmembrane; | | | | | |
| KW | Signal-anchor; Golgi stack; Multigene family. | | | | | |
| FT | DOMAIN | 1 | 7 | CYTOPLASMIC (POTENTIAL). | | |
| FT | TRANSMEM | 8 | 25 | SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL). | | |
| FT | DOMAIN | 26 | 308 | LUMENAL, CATALYTIC (POTENTIAL). | | |
| FT | CARBOHYD | 128 | 128 | N-LINKED (GLCNAC. .) (POTENTIAL). | | |
| FT | CARBOHYD | 172 | 172 | N-LINKED (GLCNAC. .) (POTENTIAL). | | |
| FT | CARBOHYD | 229 | 229 | N-LINKED (GLCNAC. .) (POTENTIAL). | | |
| SQ | SEQUENCE | 308 AA; | 35964 MW; | 789073A5178825B1 CRC64; | | |
| Query Match 75.7%; Score 1133.5; DB 1; Length 308; Best Local Similarity 73.3%; Pred. No. 1.4e+100; Matches 206; Conservative 36; Mismatches 38; Indels 1; Gaps 1; | | | | | | |
| Oy | 1 | FKEOSFYVKKD-GNFLKLDPDTDCROTPEFLVLLVTSSHKQLAAERMAIROTGWKERMYVGK | 59 | : | : | : |
| Dd | 28 | FRELPPFVKKGSHGFLQIPDDICKOKPEFLVLLVTSSHKQLAAERAIKTWGRETSGVOQG | 87 | : | : | : |
| Oy | 60 | QLKTFELLGTTSSAAETKEVDQESQRHGGIIIOKDFLDVVYNLTIKTMGMIEWHRFCPOA | 119 | : | : | : |
| Dd | 88 | QVRTFFLLGTSDSTEEMDATTTLESEQRRDIIOKFDKDAYFNLTILTKTMMGMEWHVFCEPQT | 147 | : | : | : |
| Oy | 120 | AFVMKTDSDMFINDVYLTELKKNNTRTPFTGFCLKLNEFPPIRPFPSKWFSKSEYPMDR | 179 | : | : | : |
| Dd | 148 | AIVAKTDSDMFVNNGYLTCELLLNKNTRYPTFGTYIKPHDFPIROKFNKWFYSKFEYPMDR | 207 | : | : | : |
| Oy | 180 | YPPCSGTGYYVFGDVASOVNVSKSPVIKLEDVFGVLCLERLINLEEHLHSOPTFFPG | 239 | : | : | : |
| Dd | 208 | YPPCSGTGYYVFGSDVALQTVNVSSEVPFIKLEDVFGVLCLAKLIKRPHEELUQTFFPG | 267 | : | : | : |
| Oy | 240 | GLRFSVLCFRIRIVACHFIKPRTLIDYMOALENSRGECDDCPV | 280 | : | : | : |
| Dd | 268 | GLRFSVCRFOKIVACHEFMKPODLITYMQALENSKEQDCPAV | 308 | : | : | : |
| RESULT 6 ID B3GB_HUMAN STANDARD; PRT; 372 AA. AC O9Y2A9; DT 16-OCT-2001 (Rel. 40, Created) DT 16-OCT-2001 (Rel. 40, Last sequence update) DT 15-JUN-2002 (Rel. 41, Last annotation update) DE Probable beta-1,3-galactosyltransferase 8 (EC 2.4.1.-*) (Beta-1,3-GalTase 8) (beta3Gal-T8) (b3gal-T8) (UDP-galactose-beta-N-acetylglucosamine beta-1,3-galactosyltransferase 8) (UDP-Gal:beta-GlcNAc beta-1,3-galactosyltransferase 8) (Beta-3-Gx-T8). GN B3GAUT8 OR TMEM3. OS Homo sapiens (Human). OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Euarchontia; Primates; Catarrhini; Hominoidea; Hominidae; Homo. OX NCBI_TaxID=9606; RN [1] RP SEQUENCE FROM N.A. RX MEDLINE=99173880; PubMed=10072769; RA Yokoyama-Kobayashi M., Yamaguchi T., Sekine S., Kato S.; RT "Selection of cDNAs encoding putative type II membrane proteins on the cell surface from a human full-length cDNA bank."; RL Gene 228:161-167(1999). RN RP RP REVIEW. RX MEDLINE=20047730; PubMed=10580128; RA Anado M., Almeida R., Schwientek T., Clausen H.; RT "Identification and characterization of large galactosyltransferase gene families: galactosyltransferases for all functions."; RL Biochim. Biophys. Acta 1473:35-53(1999). CC -!- PATHWAY: Glycosylation. CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (Potential). CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 31. | | | | | | |
| ----- | | | | | | |
| CC | THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute | | | | | |


```
Qy 172 KSEYPMWRYPFPCSGTGVYFGDVASQVYNNVSKSPVYIKLEDFVFGCLERINIRLEELH 231
DE : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : |
Db 294 EVVYS-GLYPPYAGGGFLYSGHLRALYHITDQVHLYPIDDVYTGMCQLGLGVPEKH 352
Qy 232 SQPTF-PFGGLRFSVCLERRIVACHFIKPRTLDDYQWALENS 272
DE : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : |
Db 353 GRTFDIEEKNKNITCSYVDLMLVHSRKPQEMIDIWSQLQA 394

RESULT 8
VP39_YEAST
ID VP39_YEAST STANDARD: PRT; 1049 AA.
AC Q07468;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Vacuolar assembly protein VPS39 (Vacuolar morphogenesis protein VAM6).
GN VPS39 OR VAM6 OR YDL077C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97269041; PubMed=9111041;
RA Nakamura N., Hirata A., Ohsumi Y., Wada Y.;
RT "Vam2/Vps41p and Vam6/Vps39p are components of a protein complex on
RT the vacuolar membranes and involved in the vacuolar assembly in the
RT yeast Saccharomyces cerevisiae."
RL J. Biol. Chem. 272:11344-11349(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Wambutt R., Wedler H., Wedler E., Scharfe M.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: REQUIRED FOR VACUOLAR ASSEMBLY.
CC -!- SUBUNIT: INTERACTS WITH VPS41/VAM2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: 274125; CAA98643.1; .
DR EMBL: D83058; BAA11758.1; .
DR SGD: S0002235; VAM6.
SQ SEQUENCE 1049 AA; 122881 MW; EE0D94E7E2F5F078 CRC64;

Query Match 7.0%; Score 105.5; DB 1; Length 1049;
Best Local Similarity 24.4%; Pred. No. 0.057;
Matches 44; Conservative 32; Mismatches 53; Indels 51; Gaps 9;

Qy 7 VYKKGDNFLKL---PD-----TDCROTFFFLVLTNS 35
DE : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : |
Db 38 VYSRDNLTUKLFQVTPDLLQNAKNDPLPPVFNFFANELSTIFAQKET---LILLST 94

Qy 36 S----HKLAERMAIRQTWGERMKVKGKQLKTFLLGTTSSAAETKEVDQESQRHGDIIQ- 91
DE : : : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : |
Db 95 NLHEVDRIIDRGINHCHWLFERSHKNEKNTYLIYSTINTAKMRVLIWEGRTYKNMMEA 154

Qy 92 -----KDFLDVY-----NLTLMGII-EWVHRFCPPAAAFVVKMTSDMFIN--VDYLT 139
DE : : : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : |
Db 155 SLSYRKRTIRSIYPGETGTITLDTLGIYHPYN-KPSLIRIEKTVKNKPKDKMISALTEL 213

RESULT 9
MTHB_HAEPH
ID MTHB_HAEPH STANDARD: PRT; 336 AA.
AC P50193;
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
```

```
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Modification methylase HphIB (EC 2.1.1.72) (Adenine-specific
DE methyltransferase HphIB) (M.HphIB) (M.HphI(A)).
GN HPHIBM OR HPHIMA.
OS Haemophilus parahaemolyticus.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=735;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 49700;
RX MEDLINE=96313243; PubMed=8759008;
RA Lubys A., Lubiene J., Kulakauskas S., Stankevicius K.,
RA Timinskas A., Janulaitis A.;
RT "Cloning and analysis of the genes encoding the type IIS restriction-
RT Modification system HphI from Haemophilus parahaemolyticus."
RL Nucleic Acids Res. 24:2760-2766(1996).
CC -!- FUNCTION: THIS METHYLASE PROTECTS THE DNA FROM CLEAVAGE BY THE
CC HPHI ENDONUCLEASE.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA adenine = S-
CC adenosyl-L-homocysteine + DNA 6-methylaminopurine.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: X85374; CAA59691.1; .
DR REBASE: 3661; M.HphIB.
DR InterPro: IPR002294; D12N6_mtfrase.
DR InterPro: IPR002052; N6_Mtase.
DR Pfam: PF02086; MethyltransfD12; 1.
DR PRINTS: PR00505; D12N6MTFRASE.
DR PROSITE: PS00092; N6_MTASE; 1.
KW Transferase; Methyltransferase; Restriction system.
SQ SEQUENCE 336 AA; 39609 MW; 3BD14C7772C4FDF CRC64;

Query Match 6.2%; Score 93.5; DB 1; Length 336;
Best Local Similarity 22.3%; Pred. No. 0.19; Mismatches 58; Indels 51; Gaps 7;
Matches 39; Conservative 27;

Qy 82 ESQRHG-DIIQKDFLDVYVYNTLK-----TMGIEWVHRFCPPAAAFVVKMTSDMFI 131
DE : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : |
Db 46 EAKKRGYQVIANDILNINYLALIVNQEIILTACDVFDFSNPKSGFMTKNYSDFVF 105

Qy 132 -----NVDTLLELLKKNRTTRFTGFLKLNFFIRQ-PFSKWFVS-----KSEYP 176
DE : : : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : |
Db 106 FKRECELDAIRANILKLNNTYKQALAFALMRAMIRKMPYSRFTISWEKVQLRDEEYS 165

Qy 177 WDRY-----PPFCSTGVYFGDVASQVYNNVSKSPVYIKLEDFV 215
DE : : : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : |
Db 166 YSKYGRRAYHNSQFPEHFRENLSYNQAVFNNGNHQAYN-----EDVF 210

RESULT 10
SS22_YEAST
ID SS22_YEAST STANDARD: PRT; 1314 AA.
AC P25390;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Serine/threonine protein kinase SSK22 (EC 2.7.-.-).
GN SSK22 OR YCR073C OR YCR73C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Ballesta J.P.G., Franco L., Hoenicka J., Jimenez A., Remacha M.,
```



```
RA Sanz E.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=95350642; PubMed=7624781;
RA Maeda T., Takekawa M., Saito H.;
RT "Activation of yeast PBS2 MAPKK by MAPKKs or by binding of an SH3-
RT containing osmosensor.";
RL Science 269:554-558(1995).
CC -!- FUNCTION: KINASE INVOLVED IN A SIGNAL TRANSDUCTION PATHWAY THAT IS
CC ACTIVATED BY CHANGES IN THE OSMOLARITY OF THE EXTRACELLULAR
CC ENVIRONMENT. ACTIVATES THE PBS2 MAP KINASE BY
CC PHOSPHORYLATION.
CC -!- SUBUNIT: INTERACTS WITH BY SSK1.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X59720; CAA42271.1; .
DR PIR; S19488; S19488.
DR HSP; P24941; B388.
DR SGD; S000669; SSK22.
DR InterPro; IPR000719; Euk_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 1034 1314 PROTEIN KINASE.
FT BIND 1040 1048 ATP (BY SIMILARITY).
FT BINDING 1063 1063 ATP (BY SIMILARITY).
FT ACT_SITE 1158 1158 BY SIMILARITY.
SQ SEQUENCE 1314 AA; 150654 MW; 8040A31FAAD56E14 CRC64;
Query Match 6.0%; Score 90.5; DB 1; Length 1314;
Best Local Similarity 19.2%; Pred. No. 2;
Matches 61; Conservative 58; Mismatches 127; Indels 71; Gaps 14;
OY 3 EOSFVYKDGNGFLKLPDTCROTTPFLVLLVTSSHKQLAERM----AIRQTW-GKERMYK 57
Db 209 EOSLTYLRRDSSDSVFEEIMKFQIPQGRKILSLDALEAILQDLNRYHSVYWPNLKMKYK 268
OY 58 GKQLKTFELGTTSSAAETKEYDOESQRHGDIIQKDFLDVYVNLTKTMGIEWVHRRCP 117
Db 269 DKPI-----TWTAETARID-----VMSWLNFKTNLTLRQELDDWNRFP 311
OY 118 QAAFMVKTDS-----DMFINVDYLTLLKKNRTTRFTG-----FLKLN 157
Db 312 ----ISSDNCQEDFDGVPQWCKMKILAEQLMKEKNIESFOKFIYPLSPWMEKLLKH 367
OY 158 EFPITROPSEKVFVSKSEYPPWDRYPFCGSGTGTVFSGDVASQV-YNVKSVYPIKLEDFV 216
Db 368 FIVRETITKWNII---KPYERLRSLLAFFVYLILKEVILTRLSVARKLNKPTMMIDQMI 424
OY 217 G--LCLERLNIRLE-----ELHSQPTF---FPGGLRFESVCLFR-RIVAC---H 255
Db 425 DDFNAFIRLSVOLKYLTKYCSNLPFFVDFTPTENTVIEAIRYLFFLLNLKLLDSSKON 484
OY 256 FIKPRTLLDYQWALENS 272
Db 485 FKAPDLLLLKYWDHLKNT 501
```

RESULT 11

```
TLRL_HUMAN
ID TLRL_HUMAN STANDARD; PRT; 786 AA.
AC Q15399; O15452; Q9UG90;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Toll-like receptor 1 precursor (Toll/Interleukin-1 receptor-like)
DE (TIL).
GN TLRL OR KIAA0012.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Erythroleukemia;
RX MEDLINE=98118556; PubMed=9435236;
RA Rock F.L., Hardiman G., Timans J.C., Kastelein R.A., Bazan J.F.;
RT "A family of human receptors structurally related to Drosophila
RT Toll.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:588-593(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=96051387; PubMed=7584026;
RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayasi Y.,
RA Sato S., Nagase T., Seki N., Ishikawa K.-I., Tabata S.;
RT "Prediction of the coding sequences of unidentified human genes. I.
RT The coding sequences of 40 new genes (K1AA0001-K1AA0040) deduced by
RT analysis of randomly sampled cDNA clones from human immature myeloid
RT cell line KG-1.";
RL DNA Res. 1:27-35(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21154917; PubMed=11230166;
RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansonge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,
RA Weves H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs.";
RL Genome Res. 11:422-435(2001).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF TIR DOMAIN.
RX MEDLINE=20531768; PubMed=11081518;
RA Xu Y., Tao X., Shen B., Horng T., Medzhitov R., Manley J.L., Tong L.;
RT "Structural basis for signal transduction by the Toll/Interleukin-1
RT receptor domains.";
RL Nature 408:111-115(2000).
CC -!- FUNCTION: Participates in the innate immune response to microbial
CC agents. Cooperates with TLR2 and modulates the response to NF-
CC microbial constituents. Acts via MyD88 and TRAF6, leading to NF-
CC kappa-B activation, cytokine secretion and the inflammatory
CC response (By similarity).
CC -!- SUBUNIT: Binds TLR2 via their respective extracellular domains.
CC Binds MyD88 via their respective TIR domains (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Plasma membrane and
CC phagosomes (By similarity).
CC -!- TISSUE SPECIFICITY: Ubiquitous. Highly expressed in spleen, ovary,
CC peripheral blood leukocytes, thymus and small intestine.
CC -!- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.
CC -!- SIMILARITY: CONTAINS 1 TIR DOMAIN.
CC -!- SIMILARITY: CONTAINS 8 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```



```
CC ----- U88540: AAC34137.1; -
DR EMBL: D13637: BAA02801.1; -
DR EMBL: AL050262: CAB43364.1; -
DR Genew: HGNC:11847; TLR1.
DR MIM: 601194; -
DR PDB: 1FYW; 22-NOV-00.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR_Cterm.
DR InterPro: IPR003591; LRR_Typ.
DR InterPro: IPR000157; TIR_domain.
DR Pfam: PF00560; LRR; 7.
DR Pfam: PF01463; LRRCT; 1.
DR Pfam: PF01582; TIR; 1.
DR PRINTS: PR00019; LEURICHRPT.
DR SMART: SM00082; LRRCT; 1.
DR SMART: SM00369; LRR_Typ; 4.
DR SMART: SM00255; TIR; 1.
DR PROSITE: PS50104; TIR; 1.
DR Receptor; Immune response; Inflammatory response; Signal;
KW Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein;
KW 3d-structure. 1 22
FT SIGNAL. 1 22
FT CHAIN 23 786
FT DOMAIN 23 580
FT TRANSMEM 581 601
FT DOMAIN 602 786
FT REPEAT 68 91
FT REPEAT 93 113
FT REPEAT 114 137
FT REPEAT 139 164
FT REPEAT 371 395
FT REPEAT 397 423
FT REPEAT 444 467
FT REPEAT 469 491
FT DOMAIN 635 779
FT CARBOHYD 51 51
FT CARBOHYD 137 137
FT CARBOHYD 163 163
FT CARBOHYD 248 248
FT CARBOHYD 330 330
FT CARBOHYD 429 429
FT CARBOHYD 578 578
FT CONFLICT 80 80
FT CONFLICT 182 182
FT CONFLICT 228 228
FT CONFLICT 248 248
FT CONFLICT 276 276
FT CONFLICT 602 602
SQ SEQUENCE 786 AA; 90290 MW; 1406C60FB3647F58 CRC64;

Query Match 6.0%; Score 89.5; DB 1; Length 786;
Best Local Similarity 22.4%; Pred. No. 1.3;
Matches 49; Conservative 38; Mismatches 79; Indels 53; Gaps 12;

QY 91 QKDFLDVYNNLTUK--TMGIEVWHRFCPO--AAFV-----MKTDSMFINVLYTE- 138
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
DB 316 QSYIYEIFSNMNIKFTVSGTRVMHMLCPKSPFLHDFSNNLLTDT-VFENCGLHTEL 374
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

QY 139 --LLKKKRTTRFTGFLKLNFFPIR-OPFSKWFVSKSEYPDRYPFCSCGTGVFSGDV 195
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
DB 375 ETLLQMNQLKE-----LSKIAETMTQMSLQQLDIDNSVSYDEKKGDCSWTKSLSLNM 430
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

QY 196 ASQVY-----NVSKSVY--YIKLEDVFVGLGLERLNIRLEELHSQP 234
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
DB 431 SSNNLTDITFRCLPPRIKVLDSLNSKIKSIPKQVVKLE-----ALQELNVAENSL- 480
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

QY 235 TFPFGGLRSVCLFRIVACHFTKPTLLDYWQALNSR 273
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
DB 481 TDLPGCGSFS--SLSLVLIIDHNSVSHPSADFFQSCQKMR 517

RESULT 12
```

```
YHV0_YEAST STANDARD; PRT; 579 AA.
AC P38848;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Hypothetical 66.1 kDa protein in IMP3-SPO12 intergenic region.
GN YHR150W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
Latrelle P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,
Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
VIII."
RL Science 265:2077-2082(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U10397; AAB68980.1; -
DR PIR: S46757; S46757.
DR SGD: S0001193; YHR150W.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 237 253
FT TRANSMEM 397 421
FT SSSEQUENCE 579 AA; 66148 MW; 2A6A73999C34C19E CRC64;

Query Match 5.9%; Score 88.5; DB 1; Length 579;
Best Local Similarity 27.4%; Pred. No. 1.1;
Matches 37; Conservative 22; Mismatches 51; Indels 25; Gaps 7;

QY 51 GKERMVKGKOL--KTF--LLGTTSSAAETKEVDQESQRHGDIIQ--KDFLDV--YVNL 102
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
DB 304 GKNEYSYGOQFFSKSFMDTLESRNQEIDEISELDKRTENTGELKQGMKVLINLRDMQMT 363
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

QY 103 LKTMMGTEWHRFCPOAAAFVAKTDSDFINVDYLTELLKKNRTRFTGFLKLNFFPIR 162
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
DB 364 SGTLLHVTEAN-----SFLRKSSS--FQN-----EECTSKRFTGFLIVFLUKIL 406
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

QY 163 QPFESKWFVSKSEYFP 177
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
DB 407 SPEVNNYSYVCISIFAW 421
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

RESULT 13
YVNF_AZOCH STANDARD; PRT; 516 AA.
AC P24423;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein in vnfD 5' region.
OS Azotobacter chroococcum mcd 1.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Azotobacter.
OX NCBI_TaxID=355;
RN 1;
```

```

RP SEQUENCE FROM N.A.
RX MEDLINE-90356423; PubMed-2388847;
RA Fallik E., Robson R.L.;
RT *Completed sequence of the region encoding the structural genes for
RL the vanadium nitrogenase of Azotobacter chroococcum.*;
RL Nucleic Acids Res. 18:4616-4616(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X51756; CAA36057.1; -
CC PIR: S14694; S14694.
CC Nitrogen fixation; Hypothetical protein.
KW SEQUENCE 516 AA; 58676 MW; ED7A7F362ED08A14 CRC64;
SQ
Query Match 5.8%; Score 86.5; DB 1; Length 516;
Best Local Similarity 24.8%; Pred. No. 1.5;
Matches 66; Conservative 30; Mismatches 97; Indels 73; Gaps 12;
QY 25 TPPLVLLVSSHQALERMAIROTGWKERMVKQKLTFFLLGTTSSAAETKEVDQ--- 81
DB 284 TPTELVDLLT-----ERRAF--TPALEQVYMATDALLSAFEVGDISTEALMFQAGYLT 334
QY 82 -ESRQHGDIQKDLVDVYVNLTKTMGIEWVHRCFQAAFPVMTKTDSDMEINVDYLTLL 140
DB 335 GQVORIG--AQPRYLRYPNLEVKASNEALLRLYLEPASVMKQTGLY-----DLL 385
QY 141 LKNKRTTRFTTGFGLKLEFFPIQRPFSKWFVSKSEYPMDRYPFPGSGTYVSGDVASQVY 200
DB 386 L-----ANDFPGQLTFTAFASIPHDVYRNPNRIARYEGY-----YASVYF 426
QY 201 NVKSVPIKLEDVFGVGLCLERLNIRLEELHSQTFPFPGGLRFSVCLFRRIVACHFIKPR 260
DB 427 S-----HFAAL---GLDVRLEDATSH-----GRIDMTV-LFNGHVLYFERKV 465
QY 261 TLLDYWQALE-----NSRGE 275
DB 466 ELTPAGQALQQLKDKGYADKYRARGE 491
RESULT 14
HSER_PIG STANDARD; PRT; 1073 AA.
AC P55204; Q29050;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Heat-stable enterotoxin receptor precursor (GC-C) (Intestinal
DE guanylate cyclase) (EC 4.6.1.2) (STA receptor).
GN GUCY2C OR GUC2C.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RX MEDLINE-95058376; PubMed-7968686;
RA Wada A., Hirayama T., Kitao S., Fujisawa J.-I., Hidaka Y.,
RA Shimonishi Y.;
RT *Pig intestinal membrane-bound receptor (guanylyl cyclase) for heat-
RT stable enterotoxin: cDNA cloning, functional expression, and
RT characterization.*;
RL Microbiol. Immunol. 38:535-541(1994).
RN [2]
RP REVISIONS TO 238 AND 509.
RA Wada A.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.

```

```

CC -!- FUNCTION: RECEPTOR FOR THE E. COLI HEAT-STABLE ENTEROTOXIN (E. COLI
CC ENTEROTOXIN MARKEDLY STIMULATES THE ACCUMULATION OF CGMP IN
CC MAMMALIAN CELLS EXPRESSING GC-C). ALSO ACTIVATED BY THE ENDOGENOUS
CC PEPTIDE GUANYLIN (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC FAMILY.
CC -!- SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D17513; BAA04465.1; -
CC HSSP: Q02846; IAWL.
DR InterPro: IPR001828; ANF_receptor.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR001054; G_cyclase.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF00211; guanylate_cyc; 1.
DR Pfam: PF01094; ANF_receptor; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00044; CYC; 1.
DR PROSITE: PS00452; GUANYLATE_CYCLASES_1; 1.
DR PROSITE: PS00125; GUANYLATE_CYCLASES_2; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Receptor; Transmembrane; Glycoprotein; Lyase; cGMP synthesis; Signal.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 1073 HEAT-STABLE ENTEROTOXIN RECEPTOR.
FT DOMAIN 24 433 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 434 454 POTENTIAL.
FT DOMAIN 455 1073 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 489 749 PROTEIN KINASE-LIKE.
FT DOMAIN 824 954 GUANYLATE CYCLASE.
FT CARBOHYD 32 32 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 43 43 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 79 79 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 284 284 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 307 307 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 357 357 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1073 AA; 123219 MW; E6F4FC1327BA9F43 CRC64;
Query Match 5.7%; Score 85; DB 1; Length 1073;
Best Local Similarity 23.1%; Pred. No. 5.2;
Matches 27; Conservative 15; Mismatches 43; Indels 32; Gaps 3;
QY 95 LDVYVNLTKTMGMIEWVHRCFQAAFPVMTKTDSDM-----FINVDY---LTEL 141
DB 103 LDLLRTISSEKRMGCVLLGPSCVSTYTFQYLDLTDLNYPMSAGSFGLSCDYKETLRLMS 162
QY 142 KKNRTTRFTTGFGLKLEFFPIQRPFSKWFVSKSEYPMDRYPFPGSGTYVSGDVASQ 198
DB 163 PARKLMYFLVDVFWVNNFP-----FKPFSWNTATYVFNKSTSE 200
RESULT 15
ETF2_MXXVL STANDARD; PRT; 711 AA.
AC Q908K4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Early transcription factor 82 kDa subunit (VETF large subunit).
GN M096L.
OS Myxoma virus (strain Lausanne).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

```

```

OC Leporipoxvirus.
OX NCBI_TaxID=31530;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20032073; PubMed=10562494;
RA Cameron C., Hota-Mitchell S., Chen L., Barrett J., Cao J.X.,
RA Macaulay C., Willer D., Evans D., McFadden G.;
RT "The complete DNA sequence of myxoma virus."
RL Virology 264:298-318(1999).
CC -!- FUNCTION: ACTS WITH RNA POLYMERASE TO INITIATE TRANSCRIPTION FROM
CC EARLY GENE PROMOTERS. A DNA-DEPENDENT ATPASE ACTIVITY IS
CC ASSOCIATED WITH VETF.
CC -!- SUBUNIT: HETERODIMER OF A 70 kDa AND A 82 kDa SUBUNIT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF170726; AAF14984.1;
KW Transcription regulation; Activator.
SQ SEQUENCE 711 AA; 82687 MW; F69EC4F63B087E61 CRC64;

Query Match          5.6%; Score 84.5; DB 1; Length 711;
Best Local Similarity 20.1%; Pred. No. 3.5;
Matches 64; Conservative 54; Mismatches 115; Indels 85; Gaps 16;

QY 4 QSFVYKKGDFLKLPTDCRQTPPELVLLVTSKQLAERMAIRQWGMKRMVKGKQLKT 63
Db :||| | | | | | | | | | | | | | | | | | | | | | | | | | |
206 KSLLYKSDYFINTLGNKFIPTDERLNLTVWDS-----DGVVTFSSDGDITMINNVKLET 260
QY 64 FELLGTTSSAAETKEVDQESQR-HGDIQKDFLDVYVNYLTLMGMGIEWVHRCFQAAFV 122
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
261 SLL-----TDIDQMERIKGDVTKYIFLST--PITSRIKLNIE-----TSFIEV 302
QY 123 MKTSDMFINVDTYLTLLKKNRTT-----RFFTGFLKL---NEFFPIRQPFQSKW 168
Db :||| | | | | | | | | | | | | | | | | | | | | | | | | | |
303 ETATNNILLSADKRISIIILAKNHSIKVKNYIPNIEKYFT-FLVIAINSFMENNIOQASDF 361
QY 169 FVSKSEYPNDRYPFCSGTGY-----VFSGDVASQVYNYKSVPIKLEDFVFG----- 217
Db :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
362 TKVETVY-WSR---ICQNTKNKHKRPVIVSSLDDEKVKSDNFYKSTKEVFINSNGIMF 417
QY 218 LCLERLN-----IRLEEL-----HSQPTFFPGGLRFVCLFRRIIVACH 255
Db ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
418 SCLDPLNKYNYVGFSLIFYLQKMCIPCCFLKNQSHTE-----TFSSCVYQKEIASD 469
QY 256 FIKPRTLDDYQALENSR 273
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
470 VINP-YILNFGKVVTKSK 486

Search completed: April 4, 2003, 14:44:27
Job time : 14 secs
```


G86397
protein T7N9.18 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: G86397
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzbarg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: G86397
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-657 <STO>
A:Cross-references: GB:AE005172; NID:g8778858; PIDN:AAF79857.1; GSPDB:GN00141
C:Genetics:
A:Gene: T7N9.18
A:Map position: 1

Query Match 17.8%; Score 267; DB 2; Length 657;
Best Local Similarity 27.5%; Pred. No. 1.3e-16;
Matches 67; Conservative 51; Mismatches 102; Indels 24; Gaps 7;

QY 43 RMAIOTWGERMVKGKOLKTFELLGTTSSAAETKEVDQESORHGDIIQKDFLDVYNYLT 102
Db 425 RMAVRKSMQOKLRSKSVARFVALHARKEVNDLKAEYFGDIVVYMDHYDLVV 484
QY 103 LKTM---MGIEWVHRFCPOAAEFVMTKDSOMFINVDYLTLLKKNRTTFFTGFLKNE 158
Db 485 LKTVAICEYGVNV-----AAKVMKCDDDTEVRVDAVIOAEAKVKGRESLIYGINFNH 539
QY 159 FPIQPFKSWFVSKSEYPWDRYPPFCSTGYVFGSDVASOVYN--VSKSVPIYKLEDDVFV 216
Db 540 KPLRT--GKAVTFEMPEEYPPYANGPGYIISYDAKFIYDDFEOKRLRFKMEVDYSM 597
QY 217 GLCLERINIRLELHSDPTFFPGGLRFSVCLFRRI---VACHIKPRTLLDYWOALENSR 273
Db 598 GMVVERFN-----ETRPVAVVSLRF--CFQGCIEDYFTHAQSPQMCMDKLRIG 649
QY 274 GEDC 277
Db 650 KPOC 653

RESULT 3
T10648
hypothetical protein T13K14.220 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 15-Oct-1999
C:Accession: T10648
R:Devan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lemcke
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z16991
A:Accession: T10648
A:Molecule type: DNA
A:Residues: 1-739 <BEV>
A:Cross-references: EMBL:AL080282; GSPDB:GN00602; ATSP:T13K14.220
A:Experimental source: cultivar Columbia; BAC clone T13K14
C:Genetics:
A:Gene: ATSP:T13K14.220
A:Map position: 4
A:Introns: 59/3; 86/1; 438/3; 535/3; 578/3; 634/3; 670/3

Query Match 17.7%; Score 264.5; DB 2; Length 739;
Best Local Similarity 27.6%; Pred. No. 2.6e-16;
Matches 71; Conservative 48; Mismatches 123; Indels 15; Gaps 6;

QY 25 TPPELVLTSSHKOLAERMAIROTWCKERMVKGKOLKTFELLGTTSSAAETKEVDQESQ 84
Db 491 TPPELVLTSSHKOLAERMAIROTWCKERMVKGKOLKTFELLGTTSSAAETKEVDQESQ 550
QY 85 RHGDIIOKDFLDVYNYLTLLKTMGMGIEWVHRFCPOAAEFVMTKDSOMFINVDYLTLLKKN 144
Db 551 YFGDIVLTPMDRYELVVLTKTIAICERGNVTAP---YIMKCDDDTFFIRVESILKQIDGVS 607
QY 145 RTRFFTGFLKNEFFPIQPFKSWFVSKSEYPWDRYPPFCSTGYVFGSDVASOV--YNV 202
Db 608 PEKSLYNGNLNLRHRLPT--GKVTWTEWPEAVPPYANGPGYIISNIAKYIVSQNS 665
QY 203 SKSVPIYKLEDDVFVGLCLERINIRLELHSDPTFFPGGLRFSV--CLFPRIVACHFIKPR 260
Db 666 RHKLRLFKMEDVSMGLWVEQFNASM-----QPVEYSHSWKFCQIGCTLLNYTA-HYQSPS 719
QY 261 TLLDYWOALENSRGDC 277
Db 720 QMMCLWDNLLKGRPQC 736

RESULT 4
D96777

hypothetical protein F25A4.23 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96777
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Al-
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kin-
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzbarg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tall-
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96777
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-642 <STO>
A:Cross-references: GB:AE005173; NID:g5882743; PIDN:AAD55296.1; GSPDB:GN00141
C:Genetics:
A:Gene: F25A4.23
A:Map position: 1

Query Match 16.3%; Score 243.5; DB 2; Length 642;
Best Local Similarity 24.9%; Pred. No. 1.9e-14;
Matches 65; Conservative 51; Mismatches 116; Indels 29; Gaps 8;

QY 28 FLVLVTSSHKOLAERMAIROTWCKERMVKGKOLKTFELLGTTSSAAETKEVDQESORHG 87
Db 396 FIGLSAGNH--FSEMAVRKSMQHVLITSAKVAVRFFVALHGRKEVNVNELKAEYFG 453
QY 88 DIQKDFLDVYNYLTLLKTMGMGIEWVHRFCPOAA-----FVMTKDSOMFINVDYLTLL 141
Db 454 DIVLVPYMDSYDLVLKTV-----AICEHALASAKYIMKCDDDTFFVKGAVINEVK 506
QY 142 KKNRTTFFTGFLKNEFFPIQPFKSWFVSKSEYPWDRYPPFCSTGYVFGSDVASOVYN 201
Db 507 KVEGRSLYIGNMNYHKPLRG--GKAVTYEEMPEEDYPPYANGPGYVLLSSDIARFIVD 564
QY 202 VSK--SVPIYKLEDDVFVGLCLERINIRLELHSDPTFFPGGLRFSVCLF---RRIVACHF 256
Db 565 KFERHLRLFKMEDVSMGMVHEF-----KNTTNPVDYRHSLRF--CFQGCVENYTAHY 617
QY 257 IKPRTLLDYWOALENSRGDC 277
Db 618 QSPQMCICLWDKLLRQNKPEC 638

RESULT 5

T20203
hypothetical protein C54C8.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T20203
R:Dobson, R.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19237
A:Accession: T20203
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-262 <N1>
A:Cross-references: EMBL:Z83102; PIDN:CAB05465.1; GSPDB:GN00019; CESP:C54C8.3
A:Experimental source: clone C54C8
C:Genetics:
A:Gene: CESP:C54C8.3
A:Map position: 1
A:Introns: 19/2; 117/3; 184/2; 226/3

Query Match 15.6%; Score 233.5; DB 2; Length 262;
Best Local Similarity 27.7%; Pred. No. 5e-14;
Matches 67; Conservative 46; Mismatches 100; Indels 29; Gaps

QY 13 NFKLPDTCROTTPPEFLVLIVTSSHKQLAERMAIQTW---GKERVMVGKQLKTFLLCT 69
Db : : | | : : : : | | : : | | : : | : : | : : | : : |
22 SWLYPTEITSEKDILIIIVASRTDSYARRNILROTWMKANSEIVANGRMKPLFLVGL 81
QY 70 TSSAAETKE-VDOESQRHGDIQOKPLDVVYNLTTLTKMGIEMWHRFCPOAAAFVMTDSD 128
Db : : | | : : : : | | : : | | : : : : : : : : : : : : : : : :
82 TPGEYMKKMVMOEAKLYGDIIVDNDNYEELTYKSAILLYGYSKAPRVQMIGKIDED 141
QY 129 MFINVDYLTELLKKNRTRFTFGFLKLNEFPF-----RQPSKWFVSKSEYP 176
Db : : | | | | : : : : | : : : : : : : : : : : : : : : : :
142 VMFFPKDLTEL-----YDQGFDPLRIYGLKMQSGANIFROKTHRWYVPESYS 192
QY 177 WDRYPPFCSCGTGVFSGDVASOVYNNVSKVPYIKLEDVFV-CLCLERLNIRLEELHSOPT 235
Db : : | | : : | :
193 CSRFPEVVSGLYKVWTWEAAQQLIISKTRYDFIQVEDVFLTGLAEDLGISVKNL---PE 249
QY 236 FF 237
Db : :
250 FY 251

RESULT 6
T20446
hypothetical protein E03H4.11 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
C:Accession: T20446
R:Dobson, R.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19276
A:Accession: T20446
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-384 <N1>
A:Cross-references: EMBL:Z81492; PIDN:CAB04032.1; GSPDB:GN00019; CESP:E03H4.11
A:Experimental source: clone E03H4
C:Genetics:
A:Gene: CESP:E03H4.11
A:Map position: 1
A:Introns: 46/1; 76/1; 196/3; 263/2; 305/3; 335/3
C:Superfamily: Caenorhabditis elegans hypothetical protein T09F5.1

Query Match 14.9%; Score 223.5; DB 2; Length 384;
Best Local Similarity 27.0%; Pred. No. 7.1e-13;
Matches 67; Conservative 44; Mismatches 104; Indels 33; Gaps

QY 9 KKGDNFLKLPDTCROTTPPEFLVLIVTSSHKQLAERMAIQTW---GKERVMVGKQLKTF 65
Db : : | : : | | : : : : | | : : | | : : | : : | : : | : : |
97 QKNYTWLYLPEIEEASQEKKDILMIIVASTDSYARRNIMROTWMKANSDSEIVANGRMKPLF 156

```
A:Molecule type: DNA
A:Residues: 1-357 <WIL>
A:Cross-references: EMBL:AL009246; PIDN:CAA15839.1; GSPDB:GN00019; CESP:C47F8.6
A:Experimental source: clone C47F8
C:Genetics:
A:Gene: CESP:C47F8.6
A:Map position: 1
A:Introns: 19/1; 52/1; 172/3; 239/2; 281/3; 311/3
C:Superfamily: Caenorhabditis elegans hypothetical protein T09F5.1

Query Match      13.4%; Score 200; DB 2; Length 357;
Best Local Similarity 25.7%; Pred. No. 9.8e-11;
Matches 59; Conservative 48; Mismatches 115; Indels 8; Gaps 4;

Qy 9 KKGNFLLPDTDCROTTPPELVLLVTSSSHQKLAERMAIROTW---GRERVMVKGKQLATFF 65
Db 73 QKNYTWLYLPTIETSKETDILMIVASRTDSYARNRMITQMTWMMKNSSEIVANGRMKSLF 132

Qy 66 LLGTTSSAAETKE-VDOESORHGDIIOKDFLDVYVYNTLTKTMGIEWVHRFCPOAAVFMK 124
Db 133 LVGLAPADYKVKVMQEAQKLYGDIIVDDMTYEELIYKSLMIFLFGVSKAPOYKLIIGK 192

Qy 125 TDSDMFINVDYLTLL---LKKNTTRFFGFLKNEFPPIROPFSKWFVSKSEYPWDRYP 181
Db 193 IDEIMFPDKLMALYEQGIIDSTPVSGLYGLVIPAGRDIFRDKTNRWYVPESAYSCSQYP 252

Qy 182 PFCSGTGVVSGDVASQVYVNSKVPYIKLEDVVFV-GLCLERLNRLEEL 230
Db 253 AYLSGMYMATREAAQMLLKSKTKRDFIQVEDVILVTGILAEGLGPIRDM 302

RESULT 9
T20031
hypothetical protein C47F8.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C:Accession: T20031; T22806
R:McLay, K.
submitted to the EMBL Data Library, December 1997
A:Reference number: Z19212
A:Accession: T20031
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-345 <WIL>
A:Cross-references: EMBL:AL009246; PIDN:CAA15841.1; GSPDB:GN00019; CESP:C47F8.3
A:Experimental source: clone C47F8
R:Kershaw, J.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19617
A:Accession: T22806
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-345 <WIL>
A:Cross-references: EMBL:Z81553; PIDN:CAB04502.1; GSPDB:GN00019; CESP:C47F8.3
A:Experimental source: clone F56H6
C:Genetics:
A:Gene: CESP:C47F8.3
A:Map position: 1
A:Introns: 23/1; 49/1; 169/3; 236/2; 278/3; 308/3
C:Superfamily: Caenorhabditis elegans hypothetical protein T09F5.1

Query Match      13.3%; Score 199.5; DB 2; Length 345;
Best Local Similarity 27.6%; Pred. No. 1e-10;
Matches 60; Conservative 39; Mismatches 107; Indels 11; Gaps 7;

Qy 29 LVLLVTSSSHQKLAERMAIROTW-GKE--RMVKGKOLKLTFFLLGTTSSAAETKE-VDOESQ 84
Db 90 ILMVLVSKTKNFARNVLRSTNRNKNSEMMKSGRMHALFFVGLVPGDQNLKLVLEAE 149

Qy 85 RHGDIQKDFLDVYVYNTLTKTMGIEWVHRFCPOAAVFMKTDSDMFINVDYLTLLKK- 143
Db 150 IHGDVHVVDLETDYDNLNPKTLALLYCTSKASQFKIIGKIDDDVNFPPDQLPLMDRNF 209

A:Molecule type: DNA
A:Residues: 1-357 <WIL>
A:Cross-references: EMBL:AL009246; PIDN:CAA15839.1; GSPDB:GN00019; CESP:C47F8.6
A:Experimental source: clone C47F8
C:Genetics:
A:Gene: CESP:C47F8.6
A:Map position: 1
A:Introns: 19/1; 52/1; 172/3; 239/2; 281/3; 311/3
C:Superfamily: Caenorhabditis elegans hypothetical protein T09F5.1

Query Match      13.4%; Score 200; DB 2; Length 357;
Best Local Similarity 25.7%; Pred. No. 9.8e-11;
Matches 59; Conservative 48; Mismatches 115; Indels 8; Gaps 4;

Qy 9 KKGNFLLPDTDCROTTPPELVLLVTSSSHQKLAERMAIROTW---GRERVMVKGKQLATFF 65
Db 73 QKNYTWLYLPTIETSKETDILMIVASRTDSYARNRMITQMTWMMKNSSEIVANGRMKSLF 132

Qy 66 LLGTTSSAAETKE-VDOESORHGDIIOKDFLDVYVYNTLTKTMGIEWVHRFCPOAAVFMK 124
Db 133 LVGLAPADYKVKVMQEAQKLYGDIIVDDMTYEELIYKSLMIFLFGVSKAPOYKLIIGK 192

Qy 125 TDSDMFINVDYLTLL---LKKNTTRFFGFLKNEFPPIROPFSKWFVSKSEYPWDRYP 181
Db 193 IDEIMFPDKLMALYEQGIIDSTPVSGLYGLVIPAGRDIFRDKTNRWYVPESAYSCSQYP 252

Qy 182 PFCSGTGVVSGDVASQVYVNSKVPYIKLEDVVFV-GLCLERLNRLEEL 230
Db 253 AYLSGMYMATREAAQMLLKSKTKRDFIQVEDVILVTGILAEGLGPIRDM 302

RESULT 10
T24925
hypothetical protein T15D6.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
C:Accession: T24925
R:Dobson, R.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19956
A:Accession: T24925
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-376 <WIL>
A:Cross-references: EMBL:Z83125; PIDN:CAB05616.1; GSPDB:GN00019; CESP:T15D6.5
A:Experimental source: clone T15D6
C:Genetics:
A:Gene: CESP:T15D6.5
A:Map position: 1
A:Introns: 59/1; 87/1; 207/3; 274/2; 316/3; 339/3
C:Superfamily: Caenorhabditis elegans hypothetical protein T09F5.1

Query Match      13.3%; Score 199; DB 2; Length 376;
Best Local Similarity 23.5%; Pred. No. 1.3e-10;
Matches 64; Conservative 60; Mismatches 120; Indels 28; Gaps 10;

Qy 9 KKGNFLLPDTDCROTTPPE---LVLLVTSSSHQKLAERMAIROTW---GKERVMVKGKQLK 62
Db 108 QKSQKRWLHPKI---KNPTWNRDILMIVASRPGSVSRKVLKRTWMNKANSKIIRNRWQ 164

Qy 63 TFFLLGTTSSAAE-TKEVDOESORHGDIIOKDFLDVYVYNTLTKTMGIEWVHRFCPOAA 121
Db 165 VLFLVGMVAGDRDLKAVKKEAESFGDIIVNMNLEDYDNLNPKVLSLLLYGNKASDFKI 224

Qy 122 VMKTSDFINVDYLTLLKK--NRTTFFTGFLKL-NEFPIROPFSKWFVSKSEYPW 178
Db 225 IKGIDDDVFFDRDLTLLDENVIDSSYIYGLSDDELVYVNRNPKWYVETAYNCT 284

Qy 179 RYPPFCSGTGVVSGDVASQVYVNSKVPYIKLEDVVFV-GLCLERLNRLEELHQSPTFF 237
Db 285 KYPVYALGPFYLIITNKAANLIVENSFRQNFMTVEDALIAAGIAGELGI---QRHSLPMNT 341

Qy 238 PGGLRFSVCLFRIRIVACHFIK--PRTLDDYQW 267
Db 342 DG-----KKILSWHMSKRSRQFLDFYQ 364

RESULT 11
T20028
hypothetical protein C47F8.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C:Accession: T20028
R:McLay, K.
submitted to the EMBL Data Library, December 1997
A:Reference number: Z19212
A:Accession: T20028
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-332 <WIL>
A:Cross-references: EMBL:AL009246; PIDN:CAA15838.1; GSPDB:GN00019; CESP:C47F8.5
A:Experimental source: clone C47F8
C:Genetics:
A:Gene: CESP:C47F8.5
A:Map position: 1
```

A; Introns: 26/1; 146/3; 213/2; 255/3; 285/4
C; Superfamily: Caenorhabditis elegans hypothetical protein T09F5.1

Query Match 12.8% Score 192; DB 2; Length 332;
Best Local Similarity 25.7%; Pred. No. 4.9e-10;
Matches 59; Conservative 45; Mismatches 118; Indels 8; Gaps

Qy 9 KKDGNFLKLPDTCRPPFLVLLVTSSHKOLAERMAIRQTW--GKERVMYKGOLKTFF 65
Db :|: ||||| |:::|| |:::|| |:::|| |:::||
47 QKTYKMVLVPETDIVLRSFDILMVASSTDFARNVLRKTMNNKNYSSEIVRDGRMKALF 106
Qy 66 LLGTSSAAATKE-VDOESQRHGDIIOKDFLDVYYNLTKTMMGIEMVHRFCPOAFAFMK 124
Db |:|: ||||| |:::|| |:::|| |:::|| |:::||
107 LVGMVSERYRRVRIMBEEAKLYGDWVIIDEDTVDDLFKSLSLLLYAVSKAPEFKVICK 166
Qy 125 TDSDMFINVDTLTLLKK---NRTTRFTFGFLKLNEFPPIQPSPSKWSKSEIYPWDRIYP 181
Db |:|: ||||| |:::|| |:::|| |:::|| |:::||
167 IDEDMVFDPDKLIPLLDGKVTDPAAPFYGGLLKEGEPIVKKHDAHWTPDYAYNCTGYP 226
Qy 182 PFCSGTGYYVGDSQVYNYSKSVPIKLED-VFVGLCLERLNIRLEEL 230
Db |:|: ||||| |:::|| |:::|| |:::|| |:::||
227 AYAGPFYLATRKAALKLVKTFEQNFMTVEDSLITGITLANDLGIPRNKL 276

RESULT 12
T20876
hypothetical protein F14B6.4 - Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
C; Accession: T20876
R; White, S.
submitted to the EMBL Data Library, November 1996
A; Reference number: Z19338
A; Accession: T20876
A; Status: preliminary; translated from GB/EMBL/DBJ
A; Molecule type: DNA
A; Residues: 1-376 <WIL>
A; Cross-references: EMBL:Z81502; PIDN:CAB04107.1; GSPDB:GN00019; CESP:F14B6.4
A; Experimental source: clone F14B6
C; Genetics:
A; Gene: CESP:F14B6.4
A; Map position: 1
A; Introns: 31/1; 69/1; 189/3; 256/2; 298/3; 328/3
C; Superfamily: Caenorhabditis elegans hypothetical protein T09F5.1

| | | | | | | | |
|-----------------------|--------|------------------------------------|------------|------------|---------|----------|------|
| Query Match | 12.74; | Score | 189.5; | DB | 2; | Length | 376; |
| Best Local Similarity | 23.6%; | Pred. No. | 9.9e-10; | | | | |
| Matches | 68; | Conservative | 57; | Mismatches | 138; | Indels | 25; |
| Gaps | 10; | | | | | | |
| Qy | 9 | KKQGNFLKLPTDCRQT-----PPF----- | LVLLVTS | SHKQLAER | MAITQW | ---GK | 52 |
| Db | 76 | KDGSSTTISFADLOKTHSNLWLYLPKPFONMSSE | ITLMI | VASRTDS | FARENVL | RKTMNPN | 135 |
| Qy | 53 | ERMVKGQKLTFFLLGTT--SSAAETKEVDQES | QRHQI | KQDFLDV | VYNTL | KTMMGIE | 110 |
| Db | 136 | SEIKQGRKALFLVGHDTGDDSRKRVVMEAR | IYGDV | VVDLKD | TVEELP | FFKSLTLL | 195 |
| Qy | 111 | WHRFCQAAFYMKTDSDMFINVDTYLT | TELLK--NRT | TRFTT | GTFL--K | LNFFP | 167 |
| Db | 196 | YGTSKASEFLIGKTIDEMFFDPKILPL | EQNLID | PSSES | IYGM | FAEGGVYR | 255 |
| Qy | 168 | WFVSKSEYPWDRIYPPFCSGTGYVFGD | VASOV | VYVSKS | VPYKLE | DVFG | 227 |
| Db | 256 | WFVPDSTGYGDMFPYTGGLFVLTQD | AAKIL | NATK | HRIF | PIEDAL | 314 |
| Qy | 228 | EELHSQPTFFPGGLRFVCLFRFRIV | ACHFIK--P | RTL | LDY--W | QALENSR | 273 |
| Db | 315 | PRIH--LPEIYGDHSLH7DEDTRGV | LAWHTG | KNDQO | QDFLDY | FVRK | 361 |

RESULT 13
B86353
protein F2E2.6 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: B86353
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Ali-
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar-
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kl-
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marz-
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tal-
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: B86353
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-398 <STO>
A:Cross-references: GB:AE005172; NID:g9280694; PIDN:AAF86563.1; CSPDB:GN00141
C:Genetics:
A:Gene: F2E2.6
A:Map position: 1

[illegible]

RESULT 14
C96573
protein F12M16.19 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: C96573
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Al-
Chai, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar
ansen, N.F.; Hughes, B.; Hulzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kil
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marz
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tal
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: C96573
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <SPTO>
A:Cross-references: GB:AE005173; MID:g7769857; PIDN:AAF69535.1; GSPDB:GN00141
C:Genetics:
A:Gene: F12M16.19
A:Map position: 1

| | | | | |
|-----------------------|------------------|------------------|------------|-------------|
| Query Match | 10.7% | Score 160; | DB 2; | Length 353; |
| Best Local Similarity | 23.6% | Pred. No. 5e-07; | | |
| Matches 61; | Conservative 42; | Mismatches 112; | Indels 44; | Gaps 8; |


```
QY 15 LKLPDTC-RQTPPFVLVLV-----TSSHKQL-----AERMAIR 47
   ||| :| ||| | :|
Db 47 LRLTGRNCLTNPPTKTVRVVMDVAGNSGVSGEKKRHKVMGPFVGTGFGSAGRRSLR 106
   ||| :| ||| | :|
QY 48 QTW-----GKERWVGKQKLTFFLLGTSSAAETKEVDQESORHGDIQKDFLDVYVNL 101
   ||| :| ||| | :|
Db 107 KTWMPDPEGLRLEESTGLAIRFMFGTKSEKMAQLREIAEYDQFVLLDIEEYSKL 166
   ||| :| ||| | :|
QY 102 TLKTMGIEWVHRFCQAA-----FVMTKDSDMFINVDYLTLLKKNRTRFTTGFGLK 156
   ||| :| ||| | :|
Db 167 PYKTLVRVICLAFFKAAAYALYDSEFYVKADDDIYLRPDLRLLAKERSHSTYLGCLK - 225
   ||| :| ||| | :|
QY 157 NEPTI-RQPFKWFVKSEYPWDRYPFCSGTGYVFGDVASQVY---NVSKSVPIKLE 212
   ||| :| ||| | :|
Db 226 -KGPVFTDKLKWYELSHLLGKEYELHAYGPIYALSADVVASLVALKNNRHALLMFNE 284
   ||| :| ||| | :|
QY 213 DVFVGLCLERLNRLEELH 231
   ||| :| ||| | :|
Db 285 DVTIGAWMLAMNVNHHNHH 303
   ||| :| ||| | :|
```

RESULT 15

```
T22387
hypothetical protein F48F7.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T22387
R:Coles, L.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z19559
A:Accession: T22387
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-253 <WIL>
A:Cross-references: EMBL:Z69661; PIDN:CAA93493.1; GSPDB:GN00028; CESP:F48F7.3
A:Experimental source: clone F48F7
C:Genetics:
A:Gene: CESP:F48F7.3
A:Map position: X
A:Introns: 13/3; 55/3; 93/2; 161/3; 219/2
```

```
Query Match          10.6%; Score 158.5; DB 2; Length 253;
Best Local Similarity 23.2%; Pred. No. 4.5e-07;
Matches 48; Conservative 45; Mismatches 87; Indels 27; Gaps 5;
```

```
QY 64 FLLGTSSAAETKEVDQESORHGDIQKDFLDVYVNLTKTMGIEWVHRFCQAAAFVM 123
   ||| :| ||| | :|
Db 19 YFVGTDIVD---DIDDEIKTYEDIIQVDINENYHNTYKAIFWINEINKCRHGPKLLL 74
   ||| :| ||| | :|
QY 124 KTDSDMFINDYLTLLKKNRTRFTTGFGLKNEFPPIQPFKWFVKSEYPWDRYPFP 183
   ||| :| ||| | :|
Db 75 KIDDDVHIDMIGLQFLIKRYSVMDDFMACRVISNGQVVRNNTSKYLSKKDYKFPNLGY 134
   ||| :| ||| | :|
QY 184 CSGTGVSFGDVASQVY-NVSKSVPIKLEDFVFGCLERLNRLEELHSQPTFFPGGLR 242
   ||| :| ||| | :|
Db 135 CGGWYFISGNLLPILHGNIEKS-QELRWDDWYVTRSL-----VG DYK 176
   ||| :| ||| | :|
QY 243 FSVCLFRIRIVACHFKPRTLDDYQWAL 269
   ||| :| ||| | :|
Db 177 ISYSLQEQ----HSISPTNVREVYSLSL 199
   ||| :| ||| | :|
```

Search completed: April 4, 2003, 14:45:30
Job time : 21 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2003, 12:08:21 ; Search time 75 Seconds
(without alignments)
550.769 Million cell updates/sec

Title: US-09-914-152-1

Perfect score: 1657

Sequence: 1 MAFPKMLMYICLLVLGALC.....TLDDYWQALENSRGDCPPV 310

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

| | |
|-----|--|
| 1: | /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.* |
| 2: | /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.* |
| 3: | /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.* |
| 4: | /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.* |
| 5: | /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.* |
| 6: | /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.* |
| 7: | /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.* |
| 8: | /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.* |
| 9: | /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.* |
| 10: | /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.* |
| 11: | /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.* |
| 12: | /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.* |
| 13: | /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.* |
| 14: | /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.* |
| 15: | /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.* |
| 16: | /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.* |
| 17: | /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.* |
| 18: | /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.* |
| 19: | /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.* |
| 20: | /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.* |
| 21: | /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.* |
| 22: | /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.* |
| 23: | /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.* |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 1657 | 100.0 | 310 | 21 | Human beta-1,3 gal |
| 2 | 1651 | 99.6 | 310 | 21 | Human beta3gal-T5 |
| 3 | 1016 | 61.3 | 192 | 23 | Human 5174397/63-2 |
| 4 | 587 | 35.4 | 326 | 15 | Beta-1,3-galactosy |
| 5 | 569 | 34.3 | 422 | 20 | Polypeptide identifi |
| 6 | 555.5 | 33.5 | 319 | 20 | Human Dendriac. H |
| 7 | 555.5 | 33.5 | 331 | 21 | Membrane-bound pro |
| 8 | 555.5 | 33.5 | 331 | 22 | Human PRO polypept |
| 9 | 555.5 | 33.5 | 331 | 22 | Human PRO1074. Ho |
| 10 | 555.5 | 33.5 | 331 | 22 | Human PRO1074 (UNQ |

| | | | | | | |
|----|-------|------|-----|----|----------|--------------------|
| 11 | 555.5 | 33.5 | 331 | 23 | AAU80224 | Human betal,3-acet |
| 12 | 550.5 | 33.2 | 331 | 22 | AA893992 | Human protein sequ |
| 13 | 497.5 | 30.0 | 193 | 23 | AAE15927 | Human 4566767/84-3 |
| 14 | 497.5 | 30.0 | 193 | 23 | AAE15928 | Mouse 2745735/84-3 |
| 15 | 480 | 29.0 | 196 | 23 | AAE15929 | Human 4502339 JBC |
| 16 | 479 | 28.9 | 196 | 23 | AAE15930 | Mouse 2745737/157- |
| 17 | 443.5 | 26.8 | 195 | 23 | AAE15933 | Mouse 2745739/84-3 |
| 18 | 433.5 | 26.2 | 195 | 23 | AAE15932 | Human 4502343 JBC |
| 19 | 416 | 25.1 | 372 | 19 | AAW64558 | Human epidermoid c |
| 20 | 416 | 25.1 | 372 | 22 | AAU29167 | Human PRO polypept |
| 21 | 416 | 25.1 | 372 | 22 | AA88404 | Human membrane or |
| 22 | 416 | 25.1 | 372 | 22 | AA849749 | Human beta 1,3-N-a |
| 23 | 416 | 25.1 | 372 | 22 | AA849750 | Human beta 1,3-N-a |
| 24 | 416 | 25.1 | 372 | 23 | AB809716 | Amino acid sequenc |
| 25 | 416 | 25.1 | 372 | 23 | AAU11272 | Human betal,3-N-ac |
| 26 | 402.5 | 24.3 | 368 | 23 | AAU11273 | Murine betal,3-N-a |
| 27 | 396 | 23.9 | 378 | 22 | AAE04752 | Human beta-1,3-gal |
| 28 | 396 | 23.9 | 378 | 23 | AAE21679 | Human galactosyltr |
| 29 | 396 | 23.9 | 378 | 23 | AB805980 | Human beta 1,3-N-a |
| 30 | 396 | 23.9 | 378 | 23 | AB807522 | Human drug metabol |
| 31 | 394 | 23.8 | 401 | 23 | AAE22148 | Human TRNFR-10 pro |
| 32 | 393 | 23.7 | 371 | 20 | AAW95171 | Protein exhibiting |
| 33 | 388 | 23.4 | 378 | 20 | AAV34171 | Human galactosyltr |
| 34 | 388 | 23.4 | 378 | 22 | AAW93908 | Human polypeptide, |
| 35 | 381.5 | 23.0 | 574 | 22 | AB863976 | Drosophila melanog |
| 36 | 375.5 | 22.7 | 334 | 22 | ABG20251 | Novel human diagno |
| 37 | 360 | 21.7 | 373 | 22 | AAW93536 | Human protein sequ |
| 38 | 360 | 21.7 | 397 | 19 | AAW80212 | Human cardiac and |
| 39 | 360 | 21.7 | 397 | 21 | AAV84683 | Amino acid sequenc |
| 40 | 360 | 21.7 | 397 | 21 | AAV69698 | Human Brainiac pro |
| 41 | 360 | 21.7 | 397 | 22 | AAG68118 | Human LiG46 polype |
| 42 | 360 | 21.7 | 413 | 22 | AA849748 | Human beta 1,3-N-a |
| 43 | 359.5 | 21.7 | 378 | 21 | AA807435 | The beta-1,3-galac |
| 44 | 359.5 | 21.7 | 402 | 21 | AA824035 | Human PRO4397 prot |
| 45 | 359.5 | 21.7 | 402 | 23 | AAE22142 | Human TRNFR-4 prot |

ALIGNMENTS

RESULT 1

AA826749

ID AA826749 standard; Protein; 310 AA.

XX AA826749;

AC AA826749;

DF 15-JAN-2001 (first entry)

XX Human beta-1,3 galactose transferase protein sequence.

XX Beta-1,3 galactose transferase; treatment; diagnosis; cancer; human;

XX digestive system.

OS Homo sapiens.

XX WO2000050608-A1.

XX 31-AUG-2000.

XX 24-FEB-2000; 2000WO-JP01070.

XX 25-FEB-1999; 99JP-0047571.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Narimatsu H, Ishiki S, Togayachi A, Sasaki K;

XX WPI; 2000-549409/50.

XX N-ESDB; AAA93875.

Beta-1,3 galactose transferase and DNA encoding it, useful for synthesis of type 1 sialyl Lewis, a carbohydrate for treatment of digestive system cancer

XX PS Claim 2; Page 97-98; 123pp; Japanese.
XX CC This invention relates to a polypeptide (I) with beta-1,3 galactose
XX CC transferase activity, or variants of (I) comprising amino acid additions,
XX CC deletions and/or substitutions. Included in the invention is DNA encoding
XX CC all or part of (I); expression vectors containing the DNA, host cells
XX CC transformed by the vectors; a method for the preparation of the
XX CC polypeptide by culture of the transformants or by expression in the milk
XX CC of a transgenic mammal, and antibodies recognising (I). The Beta-1,3
XX CC galactose transferase protein transfers galactose by beta-1,3 bonding to
XX CC N-acetylglucosamine present in a non-cyclic carbohydrate chain (such as
XX CC GlcNAc) to give Galbeta1-3GlcNAc. The protein and
XX CC DNA encoding it are useful for the treatment and diagnosis of cancer of
XX CC the digestive system. The present sequence represents the Beta-1,3
XX CC galactose transferase protein.
XX SQ Sequence 310 AA;

Query Match 100.0%; Score 1657; DB 21; Length 310;
Best Local Similarity 100.0%; Pred. No. 1.5e-178;
Matches 310; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAPPKMLMYICLLVGLGALCLYFSMYSLNPFKEQSFYKKGDFNFKLPDTCROTTPFLV 60
Db 1 MAPPKMLMYICLLVGLGALCLYFSMYSLNPFKEQSFYKKGDFNFKLPDTCROTTPFLV 60
QY 61 LLVTSSHKQLAERMAIRQTGKERMYKVGKOLKTFLLGTTSSAAETKEVDQESQRHGDII 120
Db 61 LLVTSSHKQLAERMAIRQTGKERMYKVGKOLKTFLLGTTSSAAETKEVDQESQRHGDII 120
QY 121 OKDFLDVYNYLTKTMGIEWHRFVFCPOAAFYMKTDSDMFINDVYLTLLKKKRTTRFF 180
Db 121 OKDFLDVYNYLTKTMGIEWHRFVFCPOAAFYMKTDSDMFINDVYLTLLKKKRTTRFF 180
QY 181 TGFLLKNEFFPIROPFSKFWFSKSEYPMDRYPFPCSGTGYVFGSDVASQVYNSKSVPIK 240
Db 181 TGFLLKNEFFPIROPFSKFWFSKSEYPMDRYPFPCSGTGYVFGSDVASQVYNSKSVPIK 240
QY 241 LEDVFGCLERINIRLEELHSQTPFPGLRFSVCLFRIVACHFIKPRTLDDYQWALE 300
Db 241 LEDVFGCLERINIRLEELHSQTPFPGLRFSVCLFRIVACHFIKPRTLDDYQWALE 300
QY 301 NSRGEDCPPV 310
Db 301 NSRGEDCPPV 310

RESULT 2
AAY94641
ID AAY94641 standard; Protein; 310 AA.

XX AC AAY94641;
XX DT 15-AUG-2000 (first entry)
XX DE Human beta3Gal-T5 protein sequence.
XX KW UDP-D-galactose; beta-N-acetylglucosamine beta1,3-galactosyltransferase;
XX KW Beta3Gal-T5; human; chromosome 21q22.3; galactosylation; glycoprotein;
XX KW beta1,3-galactosyl glycosylated saccharide production; glycopeptide.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Domain 13..26
XX FT /note= "Transmembrane domain"
XX PN WO200029558-A1.
XX PD 25-MAY-2000.
XX PF 11-NOV-1999; 99WO-US26807. — filed before NOV.29. 2000

XX PR 13-NOV-1998; 98DK-0001483.
XX PA (CLAU/) CLAUSEN H.
XX PI Clausen H, Amado M;
XX DR WPI; 2000-399728/34.
XX DR N-PSDB; AAA27959.
XX PT Novel nucleic acid sequence encoding human
XX PT UDP-galactose:beta-N-acetylglucosamine beta1,3-galactosyltransferase
XX PT useful for obtaining beta 1,3-galactosyl glycosylated saccharides and
XX PS glycopeptides or glycoproteins
XX PS Claim 23; Fig 1; 74pp; English.
XX CC The present invention relates to a nucleic acid sequence encoding
XX CC UDP-D-galactose:beta-N-acetylglucosamine beta1,3-galactosyltransferase
XX CC (beta3Gal-T5). Beta3 transferases add galactose to the hydroxy group at
XX CC carbon 3 of 2-acetamido-2-deoxy-D-glucose (GlcNAc). The present sequence
XX CC represents the human beta3Gal-T5 protein sequence. The beta3Gal-T5 gene
XX CC is located on human chromosome 21q22.3. Beta3Gal-T5 is a type II
XX CC transmembrane glycoprotein. The invention also relates to the beta3Gal-T5
XX CC protein sequence, a nucleic acid vector comprising the beta3Gal-T5
XX CC nucleotide sequence, a host cell comprising the vector, and a method for
XX CC the production of the beta3Gal-T5 protein from the host cells. The
XX CC methods of the invention can be used for recombinant production of
XX CC beta3Gal-T5 for use as a catalyst and for recombinant production of
XX CC peptides or proteins with appropriate galactosylation. The beta3Gal-T5
XX CC protein can be used to obtain beta1,3-galactosyl glycosylated
XX CC saccharides, glycopeptides or glycoproteins.
XX SQ Sequence 310 AA;

Query Match 99.6%; Score 1651; DB 21; Length 310;
Best Local Similarity 99.7%; Pred. No. 7.1e-178;
Matches 309; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAPPKMLMYICLLVGLGALCLYFSMYSLNPFKEQSFYKKGDFNFKLPDTCROTTPFLV 60
Db 1 MAPPKMLMYICLLVGLGALCLYFSMYSLNPFKEQSFYKKGDFNFKLPDTCROTTPFLV 60
QY 61 LLVTSSHKQLAERMAIRQTGKERMYKVGKOLKTFLLGTTSSAAETKEVDQESQRHGDII 120
Db 61 LLVTSSHKQLAERMAIRQTGKERMYKVGKOLKTFLLGTTSSAAETKEVDQESQRHGDII 120
QY 121 OKDFLDVYNYLTKTMGIEWHRFVFCPOAAFYMKTDSDMFINDVYLTLLKKKRTTRFF 180
Db 121 OKDFLDVYNYLTKTMGIEWHRFVFCPOAAFYMKTDSDMFINDVYLTLLKKKRTTRFF 180
QY 181 TGFLLKNEFFPIROPFSKFWFSKSEYPMDRYPFPCSGTGYVFGSDVASQVYNSKSVPIK 240
Db 181 TGFLLKNEFFPIROPFSKFWFSKSEYPMDRYPFPCSGTGYVFGSDVASQVYNSKSVPIK 240
QY 241 LEDVFGCLERINIRLEELHSQTPFPGLRFSVCLFRIVACHFIKPRTLDDYQWALE 300
Db 241 LEDVFGCLERINIRLEELHSQTPFPGLRFSVCLFRIVACHFIKPRTLDDYQWALE 300
QY 301 NSRGEDCPPV 310
Db 301 NSRGEDCPPV 310

RESULT 3
AAE15931
ID AAE15931 standard; Protein; 192 AA.

XX AC AAE15931;
XX DT 26-MAR-2002 (first entry)
XX DE Human 5174397/63-284 protein, member of Fringe and Brainiac family.

XX Fringe protein; Brainiac protein; glycosyltransferase; EGF; mannose;
 KW N-acetylglucosamine; fucose; epidermal growth factor; T cell leukaemia;
 KW breast cancer; stroke; dementia; leucoencephalopathy; Alagille syndrome;
 KW notch protein; EGF-module containing protein; human.
 XX Homo sapiens.
 XX WO200187321-A2.
 XX 22-NOV-2001.
 XX 21-MAY-2001; 2001WO-1B01033.
 XX 19-MAY-2000; 2000GB-0012216.
 XX (EURO-) EURO MOLECULAR BIOLOGY LAB.
 XX Cohen S, Brueckner K, Clausen H, Keck B;
 PI WPI; 2002-082940/11.
 XX Use of Fringe protein or Brainiac protein as glycosyltransferase and
 PT treatment of disease caused by epidermal growth factor-like module
 PT containing protein, such as T cell leukemia, breast cancer, stroke and
 PT dementia.
 XX Claim 7; Fig 1; 52pp; English.
 XX The invention relates to the use of a Fringe protein or a Brainiac
 CC protein or a fragment or functional equivalent of a Fringe or a
 CC Brainiac protein, as glycosyltransferase. Fringe and Brainiac
 CC proteins have been found to possess glycosyltransferase activity in
 CC transferring sugar residues onto certain proteins of biological
 CC interest, so affecting the binding of effector molecules to these
 CC proteins. Particularly they are useful for transferring a
 CC N-acetylglucosamine moiety onto a fucose or a mannose substrate,
 CC whether free or attached to a lipid, carbohydrate or protein. They
 CC act as glycosyltransferase on notch protein (EGF-module containing
 CC protein) family. They are useful in treatment of a disease caused
 CC by epidermal growth factor (EGF)-like module containing protein,
 CC such as T cell leukaemia, breast cancer, stroke, dementia, cerebral
 CC autosomal dominant arteriopathy with subcortical infarcts,
 CC leucoencephalopathy and Alagille syndrome. Fringe and Brainiac
 CC proteins are also useful for screening ligands capable of modulating
 CC the activity of the protein. The present sequence is human protein
 CC which is a member of Fringe and Brainiac protein family.
 XX SQ Sequence 192 AA;

Query Match 61.38; Score 1016; DB 23; Length 192;
 Best Local Similarity 100.0%; Pred. No. 2.5e-106;
 Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 63 VTSSHKQLAERMAIRQTWGERMVKGKLTFFLLGTTSSAAETKEVDQESORHGDIQK 122
 DB 1 VTSSHKQLAERMAIRQTWGERMVKGKLTFFLLGTTSSAAETKEVDQESORHGDIQK 60
 QY 123 DFLDVYNNLTUKTMGIEWHRCFQAAAFVMTDSMFINDYLTLLKKNRTRFTG 182
 DB 61 DFLDVYNNLTUKTMGIEWHRCFQAAAFVMTDSMFINDYLTLLKKNRTRFTG 120
 QY 183 FLKLINEPIRQPSKWFVSKSEYPDWDRYPPPCSGTGYVFGDASQVYVNSKSPYIKLE 242
 DB 121 FLKLINEPIRQPSKWFVSKSEYPDWDRYPPPCSGTGYVFGDASQVYVNSKSPYIKLE 180
 QY 243 DVFVGLCLERLN 254
 DB 181 DVFVGLCLERLN 192

RESULT 4
 AAR57433

ID AAR57433 standard; Protein; 326 AA.
 XX
 AC AAR57433;
 DT 14-MAR-1995 (first entry)
 XX
 DE Beta-1,3-galactosyltransferase.
 XX
 KW Beta-1,3-galactosyltransferase; enzyme; saccharide chain; PAMOPRWM1;
 KW KJM-1 cells.
 XX
 OS Homo sapiens.
 XX JP06181759-A.
 XX 05-JUL-1994.
 XX 16-DEC-1992; 92JP-0336436.
 XX 16-DEC-1992; 92JP-0336436.
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 XX WPI; 1994-251683/31.
 DR N-PSDB; AAR57433.
 XX
 PT Beta-galactosyl-transferase DNA and protein - useful for prodn. of
 PT saccharide chains
 XX
 PS Claim 1; Page 22-24; 47pp; Japanese.
 XX This sequence represents a beta-1,3-galactosyltransferase. This
 CC enzyme can be used to produce physiologically active saccharide
 CC chains and variants, and for improvement of saccharide chains bound
 CC to physiologically active proteins. The cDNA encoding this protein
 CC represents a fragment of the plasmid PAMOPRWM1 which was cloned in
 CC KJM-1 cells.
 XX SQ Sequence 326 AA;
 Query Match 35.48; Score 587; DB 15; Length 326;
 Best Local Similarity 38.68; Pred. No. 2e-57;
 Matches 127; Conservative 52; Mismatches 108; Indels 42; Gaps 7;
 QY 12 CLLVGLGALCLYFSMYSLN-----PFKEQSFVYKK--DGN-----FL 45
 DB 7 CLYLVTVVCASALMWLITRTPTSSYTGSKPFSHLTVARKNFTFGNIRTPINPHSFEFL 66
 QY 46 KLPTDTCROTTPPFLVLLVTSSHKQLAERMAIRQTWGERMVKGKLTFFLLGTTSSAAE 105
 DB 67 INEPNCKENIPFVLILISTTHKEFDARQAIRETWGDENNFKGIKTIATLLGKNADPVL 126
 QY 106 TKEVDQESORHGDIQKDFLDVYNNLTUKTMGIEWHRCFQAAAFVMTDSMFINDY 165
 DB 127 NQMVQESQIFHDIIVDFIDSYHNLTLKTLGMRVAVTFCSKAKYVMTDSIDFYNNMDN 186
 QY 166 LTELKLNKR--TTRFFTGFLKLNFFPIRQPSKWFVSKSEYPDWDRYPPPCSGTGYVFG 223
 DB 187 LLYLKLKSTKPRRYFTGYV-INGPIRDVSKWYMPDLPDSNYPFCSTGTGIFSA 245
 QY 224 DVASQVYVNSKSPYIKLEDFVGLCLERLNIRLEELHSQPTFFPG-----LRFSCVCLF 278
 DB 246 DVAELIYKTSLHRLLEDVYVGLCLRLGIH-----PFQNSGFNHWKMYSLCRY 297
 QY 279 RRVACHIKPRTLLDYQVALENSRGDC 307
 DB 298 RRVITVHQISPEEMHRIWDMSSKKHLRC 326
 RESULT 5
 AAY02366
 ID AAY02366 standard; Protein; 422 AA.
 XX

PR 16-JUN-1998; 98US-0089512.
 PR 16-JUN-1998; 98US-0089514.
 PR 17-JUN-1998; 98US-0089532.
 PR 17-JUN-1998; 98US-0089538.
 PR 17-JUN-1998; 98US-0089598.
 PR 17-JUN-1998; 98US-0089599.
 PR 17-JUN-1998; 98US-0089600.
 PR 17-JUN-1998; 98US-0089653.
 PR 16-JUN-1998; 98US-0089801.
 PR 18-JUN-1998; 98US-0089907.
 PR 18-JUN-1998; 98US-0089908.
 PR 19-JUN-1998; 98US-0089947.
 PR 19-JUN-1998; 98US-0089948.
 PR 19-JUN-1998; 98US-0089952.
 PR 22-JUN-1998; 98US-0090246.
 PR 22-JUN-1998; 98US-0090252.
 PR 22-JUN-1998; 98US-0090254.
 PR 23-JUN-1998; 98US-0090349.
 PR 23-JUN-1998; 98US-0090355.
 PR 24-JUN-1998; 98US-0090429.
 PR 24-JUN-1998; 98US-0090431.
 PR 24-JUN-1998; 98US-0090435.
 PR 24-JUN-1998; 98US-0090444.
 PR 24-JUN-1998; 98US-0090445.
 PR 24-JUN-1998; 98US-0090461.
 PR 24-JUN-1998; 98US-0090472.
 PR 24-JUN-1998; 98US-0090535.
 PR 24-JUN-1998; 98US-0090538.
 PR 24-JUN-1998; 98US-0090540.
 PR 24-JUN-1998; 98US-0090557.
 PR 25-JUN-1998; 98US-0090676.
 PR 25-JUN-1998; 98US-0090678.
 PR 25-JUN-1998; 98US-0090688.
 PR 25-JUN-1998; 98US-0090690.
 PR 25-JUN-1998; 98US-0090691.
 PR 25-JUN-1998; 98US-0090694.
 PR 25-JUN-1998; 98US-0090695.
 PR 25-JUN-1998; 98US-0090696.
 PR 26-JUN-1998; 98US-0090862.
 PR 26-JUN-1998; 98US-0090863.
 PR 01-JUL-1998; 98US-0091358.
 PR 01-JUL-1998; 98US-0091360.
 PR 01-JUL-1998; 98US-0091544.
 PR 02-JUL-1998; 98US-0091478.
 PR 02-JUL-1998; 98US-0091486.
 PR 02-JUL-1998; 98US-0091519.
 PR 02-JUL-1998; 98US-0091626.
 PR 02-JUL-1998; 98US-0091628.
 PR 02-JUL-1998; 98US-0091633.
 PR 02-JUL-1998; 98US-0091646.
 PR 02-JUL-1998; 98US-0091573.
 PR 07-JUL-1998; 98US-0091978.
 PR 07-JUL-1998; 98US-0091982.
 PR 09-JUL-1998; 98US-0092182.
 PR 10-JUL-1998; 98US-0092472.
 PR 20-JUL-1998; 98US-0093339.
 PR 30-JUL-1998; 98US-0094651.
 PR 04-AUG-1998; 98US-0095282.
 PR 04-AUG-1998; 98US-0095285.
 PR 04-AUG-1998; 98US-0095301.
 PR 04-AUG-1998; 98US-0095302.
 PR 04-AUG-1998; 98US-0095318.
 PR 04-AUG-1998; 98US-0095321.
 PR 04-AUG-1998; 98US-0095325.
 PR 10-AUG-1998; 98US-0095916.
 PR 10-AUG-1998; 98US-0095929.
 PR 10-AUG-1998; 98US-0096012.
 PR 11-AUG-1998; 98US-0096143.
 PR 11-AUG-1998; 98US-0096146.
 PR 12-AUG-1998; 98US-0096329.
 PR 17-AUG-1998; 98US-0096757.
 PR 17-AUG-1998; 98US-0096766.
 PR 17-AUG-1998; 98US-0096768.
 PR 17-AUG-1998; 98US-0096773.
 PR 17-AUG-1998; 98US-0096791.
 PR 17-AUG-1998; 98US-0096867.
 PR 17-AUG-1998; 98US-0096891.
 PR 17-AUG-1998; 98US-0096894.
 PR 17-AUG-1998; 98US-0096895.
 PR 17-AUG-1998; 98US-0096897.
 PR 18-AUG-1998; 98US-0096949.
 PR 18-AUG-1998; 98US-0096950.
 PR 18-AUG-1998; 98US-0096959.
 PR 18-AUG-1998; 98US-0096960.
 PR 18-AUG-1998; 98US-0097022.
 PR 19-AUG-1998; 98US-0097141.
 PR 20-AUG-1998; 98US-0097218.
 PR 24-AUG-1998; 98US-0097661.
 PR 26-AUG-1998; 98US-0097951.
 PR 26-AUG-1998; 98US-0097952.
 PR 26-AUG-1998; 98US-0097954.
 PR 26-AUG-1998; 98US-0097955.
 PR 26-AUG-1998; 98US-0097971.
 PR 26-AUG-1998; 98US-0097974.
 PR 26-AUG-1998; 98US-0097978.
 PR 26-AUG-1998; 98US-0097979.
 PR 26-AUG-1998; 98US-0097986.
 PR 26-AUG-1998; 98US-0098014.
 PR 31-AUG-1998; 98US-0098525.
 PR 16-SEP-1998; 98US-0100634.
 PR 12-JAN-1999; 99US-0115565.
 (GETH) GENENTECH INC.
 Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;
 PI Wood WI, Yuan J;
 WPI: 2000-072883/06.
 N-PSDB: AAZ65022.
 Membrane-bound proteins and related nucleotide sequences -
 claim 12; Fig 137; 822pp; English.
 The invention provides membrane-bound PRO polypeptides and polynucleotides encoding them. The PRO sequences of the invention were identified based on extracellular domain homology screening. The PRO sequences have homology with proteins including LDL receptors, TIE ligands and various enzymes. The membrane-bound proteins and receptor molecules are useful as pharmaceutical and diagnostic agents. Receptor immunoadhesins, for instance, can be used as therapeutic agents to block receptor-ligand interactions. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. The PRO encoding sequences are useful as hybridization probes, in chromosome and gene mapping and in the generation of antisense RNA and DNA. PRO nucleic acid sequences will also be useful for the preparation of PRO polypeptides, especially by recombinant techniques.
 Sequence 331 AA;
 Query Match 33.5%; Score 555.5; DB 21; Length 331;
 Best Local Similarity 36.2%; Pred. No. 7.4e-54;
 Matches 112; Conservative 57; Mismatches 125; Indels 15; Gaps 2
 QY 6 MRLMYICLLVLGALCLYFSMYSLNP-----FKEOSFVYKKGDFLKLPTDCR 53
 Db 15 LRSLLKWSLLLSLLSFFVMYLSLPHYNVTVRYNMYFYEFYEQDFHTLREHNSCS 74
 QY 54 QTPFPLVLLVTSKHLAERMAIRQTWGERMYKQKLTFFLLGTTSSAAE---TKEYD 110
 Db 75 HQNPFVLVTSHPDVYKARQIRVTWGEKKSNWGYEVLTFFLLGQAEKEDMLALSLE 134
 QY 111 QESQRHGDIIQKDFLDVYNYNLTKTWGIEWHRFCFAAFYKMTOSDFINVDYTELL 170
 Db 135 DEHLLYGDIIRQDFLDTYNLTKTMAFNRVTFPCPNARYKMTDTDFINTGLVKYL 194

```

Qy 171 LKKNRTTFFTGFLKLNFFPIQPFKSWFVSKSEYPMWRPFCSGTGYVFGDVASQVY 230
Db 195 LNLHSEKFFTGYPIDNYSYRGFKTHISYQYFVFKVFPYCSGLGYMSRDLVPIY 254
Qy 231 NVSKSVPIKLEDFVGLCLERLNIRLELHLSQPTFFPGGLRFSVCLFRRIVACHFIKPR 290
Db 255 EMGHVKPIKFDVYGVGICLLKLVNIHIPEDTNLFLLYRIHLVDVQLRRVIAAHGFSSK 314
Qy 291 TLLDYWQAL 299
Db 315 EIITFQVM 323

RESULT 8
AAU29104
ID AAU29104 standard; Protein: 331 AA.
AC AAU29104;
XX DT
XX 18-DEC-2001 (first entry)
XX DE
XX Human PRO polypeptide sequence #81.
XX KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
XX KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
XX KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
XX KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
XX OS Homo sapiens.
XX PN WO200168848-A2.
XX XZ
XX PD 20-SEP-2001.
XX XZ
XX PF 28-FEB-2001; 2001WO-US06520.
XX PR 01-MAR-2000; 2000WO-US05601.
XX PR 02-MAR-2000; 2000WO-US05841.
XX PR 03-MAR-2000; 2000US-187202P.
XX PR 06-MAR-2000; 2000US-186968P.
XX PR 14-MAR-2000; 2000US-189320P.
XX PR 14-MAR-2000; 2000US-189328P.
XX PR 15-MAR-2000; 2000US-189328P.
XX PR 15-MAR-2000; 2000WO-US06884.
XX PR 21-MAR-2000; 2000US-190828P.
XX PR 21-MAR-2000; 2000US-191007P.
XX PR 21-MAR-2000; 2000US-191048P.
XX PR 21-MAR-2000; 2000US-191314P.
XX PR 28-MAR-2000; 2000US-192655P.
XX PR 29-MAR-2000; 2000US-193032P.
XX PR 29-MAR-2000; 2000US-193032P.
XX PR 29-MAR-2000; 2000US-193032P.
XX PR 30-MAR-2000; 2000WO-US08439.
XX PR 04-APR-2000; 2000US-194449P.
XX PR 11-APR-2000; 2000US-194647P.
XX PR 11-APR-2000; 2000US-195975P.
XX PR 11-APR-2000; 2000US-196000P.
XX PR 11-APR-2000; 2000US-196187P.
XX PR 11-APR-2000; 2000US-196690P.
XX PR 11-APR-2000; 2000US-196820P.
XX PR 18-APR-2000; 2000US-198121P.
XX PR 18-APR-2000; 2000US-198585P.
XX PR 23-APR-2000; 2000US-199397P.
XX PR 23-APR-2000; 2000US-199550P.
XX PR 23-APR-2000; 2000US-199654P.
XX PR 03-MAY-2000; 2000US-201516P.
XX PR 17-MAY-2000; 2000WO-US13705.
XX PR 22-MAY-2000; 2000WO-US14042.
XX PR 30-MAY-2000; 2000WO-US14941.
XX PR 02-JUN-2000; 2000WO-US15284.
XX PR 05-JUN-2000; 2000US-209832P.
XX PR 28-JUL-2000; 2000WO-US20710.
XX PR 22-AUG-2000; 2000US-0644848.
XX PR 24-AUG-2000; 2000WO-US23328.

```

```

PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
XX (GETH ) GENENTECH INC.
XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
XX Pan J, Smith V, Watanabe CK, Wood W1, Zhang Z;
XX WPI; 2001-602746/68.
XX N-PSDB; AAS46005.
XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
XX presence of tumours, such as prostate and breast tumours, in mammals and
XX to screen for modulators of the compounds -
XX Claim 11; Fig 162; 774pp; English.
XX Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
XX The PRO polypeptides and their associated nucleic acids can be used to
XX detect the presence of a tumour in a mammal by comparing the level of
XX expression of a PRO polypeptide in a test sample of cells from the animal
XX and a control sample of normal cells, whereby a higher level of
XX expression in the test sample indicates the presence of a tumour in the
XX mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
XX and rabbits but are preferably human. The polypeptides can be used to
XX stimulate tumour necrosis factor (TNF) alpha release from human blood,
XX when contacted with it. A specific polypeptide can be used to stimulate
XX the proliferation or differentiation of chondrocyte cells. The PRO
XX proteins can be used to determine the presence of tumours and also
XX susceptibility to tumour development, particularly adrenal, lung, colon,
XX breast, prostate, rectal, cervical, or liver tumours, in mammalian
XX subjects. The oligonucleotide probes specific for the PRO nucleic acids
XX can be used for genetic analysis of individuals with genetic disorders.
XX SQ Sequence 331 AA;

Query Match 33.5%; Score 555.5; DB 22; Length 331;
Best Local Similarity 36.2%; Pred. No. 7.4e-54;
Matches 112; Conservative 57; Mismatches 125; Indels 15; Gaps 2;

Qy 6 MRLMYICLLVLGALCLYFSMYSINP-----FKEQSFVYKKDGNFLKLPOTDCR 53
Db 15 LRSKLWSLLLSLSPFFVMWYLSLPHYNIERNVMYFVEYPIYRQDFHTLREHSNCS 74
Qy 54 OTTPELVLLVTSSSHKQLAERMAIROTWCKERNVKGKQLKTFLLGTTSSAAE---TKEYVD 110
Db 75 HONPELVLLVTSHPDSVVKARQAIRVTWGEKSKWNGEYVLTFFLLGQEAEEKDKMLALSLE 134
Qy 111 QESQRHGDIIQKDFLDVYVNLTKTMWGIEWVHRFCPOAAFYVKTDSDFINVDYLTLL 170
Db 135 DEHLLYGDIIRODFLDYNNLTLTIMAFRWTECPNKNKYVKTDTDFVINTGNLVKYL 194
Qy 171 LKKNRTTFFTGFLKLNFFPIQPFKSWFVSKSEYPMWRPFCSGTGYVFGDVASQVY 230
Db 195 LNLHSEKFFTGYPIDNYSYRGFKTHISYQYFVFKVFPYCSGLGYMSRDLVPIY 254
Qy 231 NVSKSVPIKLEDFVGLCLERLNIRLELHLSQPTFFPGGLRFSVCLFRRIVACHFIKPR 290
Db 255 EMGHVKPIKFDVYGVGICLLKLVNIHIPEDTNLFLLYRIHLVDVQLRRVIAAHGFSSK 314
Qy 291 TLLDYWQAL 299
Db 315 EIITFQVM 323

RESULT 9
AAB87543
ID AAB87543 standard; Protein: 331 AA.
XX AC
XX AAB87543;
XX DT 15-MAY-2001 (first entry)

```


XX DE Human PRO1074.
XX KW Human; PRO protein; mapping.
XX OS Homo sapiens.
XX PN WO200116318-A2.
XX PD 08-MAR-2001.
XX PF 24-AUG-2000; 2000WO-US23328.
XX PR 01-SEP-1999; 99WO-US20111.
XX PR 15-SEP-1999; 99WO-US21090.
XX PR 07-DEC-1999; 99US-0169495.
XX PR 09-DEC-1999; 99US-0170262.
XX PR 11-JAN-2000; 2000US-0175481.
XX PR 18-FEB-2000; 2000WO-US04341.
XX PR 18-FEB-2000; 2000WO-US04342.
XX PR 22-FEB-2000; 2000WO-US04414.
XX PR 01-MAR-2000; 2000WO-US05601.
XX PR 03-MAR-2000; 2000US-0187202.
XX PR 25-APR-2000; 2000US-0199397.
XX PR 22-MAY-2000; 2000WO-US14042.
XX PR 05-JUN-2000; 2000US-0209832.
XX PA (GETH) GENENTECH INC.
XX PI Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
XX PI Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;
XX DR WPI: 2001-183260/18.
XX DR N-PSDB; AAF92075.
XX PT Eighty four nucleic acids encoding PRO polypeptides, useful in
XX PT molecular biology, including use as hybridization probes, and in
XX PT chromosome and gene mapping.
XX PS Claim 12; Fig 36; 278pp; English.
XX CC The present sequence is a human PRO polypeptide (secreted and
XX CC transmembrane). The PRO protein, and PRO agonists, PRO antagonists or
XX CC anti-PRO antibodies are useful for preparation of a medicament useful in
XX CC the treatment of a condition which is responsive to the PRO protein,
XX CC agonists, antagonists or anti-PRO antibodies. The PRO protein may also be
XX CC employed as molecular weight markers for protein electrophoresis. The PRO
XX CC coding sequence has applications in molecular biology, including use as
XX CC hybridisation probes, and in chromosome and gene mapping.
XX SQ Sequence 331 AA;
Query Match 33.5%; Score 555.5; DB 22; Length 331;
Best Local Similarity 36.2%; Pred. No. 7.4e-54;
Matches 112; Conservative 57; Mismatches 125; Indels 15; Gaps 2;
QY 6 MRLMYICLLVGLALCLYFMSYSLNP-----FKEQSFYVKKDGNFLKLPDTCR 53
DB 15 LRLSKWSLLLSLLSPFVVMYLSLPHYVNIERNVNMWFEYEPYRQDFHTLREHSNC 74
QY 54 OTPPFLVLLVTSKQLAERMAIRQTWGERMVKGKOLKTFLLGTTSSAAE---TRKVD 110
DB 75 HQNPFLVILVTSHPSDVKARQAIRVTWGERKSWNGYEVLTFFLLGQEAKEKDKMLALSLE 134
QY 111 QESQRHGDITQKDFLDVYVNLTKTMGIEWHRECFQAAAFVMTKDSDMINVDYLTLL 170
DB 135 DEHLVGDIIIRDFLDYTNLTKTMAFRWVTEFCFNAKYVMKTDVDFINTGNLVKYL 194
QY 171 LKNRRTRFTGKLNKEFFIRPFKSWFYSKSEYKSEYKSEYKSEYKSEYKSEYKSEYKSEY 230
DB 195 LNLNHEKFTGTGLDINYSYRGFYQKTHISYQYEPKFPYPPYCSGLGYTMSRDLVPR 254
QY 231 NVSKSVPIKLEDFVGLCLERLNIREEELHSQPTFFPGGLRFSVCLFRIRVACHFIKPR 290

Db 255 EMGHRVKPIKFEDVYVGICLNLKVNHIHPEDTNLFFLYRIHLDVCQLRRVIAAHGFSSK 314
QY 291 TLLDYWQAL 299
Db 315 EIITFWQVM 323
RESULT 10
AAB65208
ID AAB65208 standard; Protein; 331 AA.
XX AC AAB65208;
XX DT 02-APR-2001 (first entry)
XX DE Human PRO1074 (UNQ531) protein sequence SEQ ID NO:209.
XX KW Human; secreted and transmembrane protein; PRO; cytostatic;
XX KW cell death; cancer; chromosomal mapping; gene mapping; tissue typing;
XX KW diagnostic assay.
XX OS Homo sapiens.
XX PN WO200073454-A1.
XX PD 07-DEC-2000.
XX PF 30-MAR-2000; 2000WO-US08439.
XX PR 02-JUN-1999; 99WO-US12252.
XX PR 23-JUN-1999; 99US-0141037.
XX PR 07-JUL-1999; 99US-0143048.
XX PR 20-JUL-1999; 99US-0144758.
XX PR 26-JUL-1999; 99US-0145698.
XX PR 28-JUL-1999; 99US-0146222.
XX PR 17-AUG-1999; 99US-0149396.
XX PR 15-SEP-1999; 99WO-US21090.
XX PR 08-SEP-1999; 99US-0158663.
XX PR 30-NOV-1999; 99WO-US28313.
XX PR 01-DEC-1999; 99WO-US28301.
XX PR 16-DEC-1999; 99WO-US30095.
XX PR 20-DEC-1999; 99WO-US30911.
XX PR 05-JAN-2000; 2000WO-US00219.
XX PR 06-JAN-2000; 2000WO-US00376.
XX PR 11-FEB-2000; 2000WO-US03565.
XX PR 18-FEB-2000; 2000WO-US04341.
XX PR 22-FEB-2000; 2000WO-US04414.
XX PR 24-FEB-2000; 2000WO-US04914.
XX PR 24-FEB-2000; 2000WO-US05004.
XX PR 02-MAR-2000; 2000WO-US05841.
XX PR 15-MAR-2000; 2000WO-US06884.
XX PR 20-MAR-2000; 2000WO-US07377.
XX PA (GETH) GENENTECH INC.
XX PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
XX PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;
XX PI Grimaldi CJ, Gurney AL, Kljavin LJ, Napier MA, Pan J, Paoni NF;
XX PI ROY MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;
XX PI Zhang Z;
XX WPI: 2001-032160/04.
XX N-PSDB; AAF44168.
XX PT PRO polynucleotides used to produce polypeptides used to target
XX PT bioactive molecules such as toxins, radiolabels or antibodies, to
XX PT specific cells, to cause targeted cell death -
XX PS Claim 12; Fig 137; 935pp; English.
XX CC The present invention describes human secreted and transmembrane PRO

CC proteins. The PRO proteins have cytostatic activity. The PRO proteins
 CC can be used for targeted delivery of bioactive molecules, such as
 CC toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide
 CC sequences, and their fragments, can be used as hybridisation probes, in
 CC chromosomal and gene mapping, and in the generation of anti-sense RNA
 CC and DNA. They may also be used to produce transgenic animals which are
 CC used to develop and screen therapeutically useful reagents. The PRO
 CC nucleotide and protein sequence can be used for tissue typing and in
 CC treating cancer. Anti-PRO antibodies can be used in diagnostic assays.
 CC AAF44270 to AAF44470 represent PCR primers and hybridisation probes used
 CC in the isolation of human PRO sequences. AAF44087 to AAF44269 and
 CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein
 CC sequences given in the exemplification of the present invention.

XX Sequence 331 AA;

Query Match 33.5%; Score 555.5; DB 22; Length 331;
 Best Local Similarity 36.2%; Pred. No. 7.4e-54;
 Matches 112; Conservative 57; Mismatches 125; Indels 15; Gaps 2;

QY 6 MRLMYICLLVLCALCLYFSMYSLNP-----FKEQSFVYKKGDNFLKLPDTCR 53
 DB 15 LRSKWSLLLSLSSFFVMYLSLPHYNVIERVNMWYFYEPIYRQDFHFTLREHSNCS 74
 QY 54 QTPPELVLLVSSHKQLAERMAIROTWGKRMVKGKOLKTFELLGTTSSAAE---TKEVD 110
 DB 75 HONPELVILVTHSPSDVKARQAIRVTWGEKSWGVEVLTFLLGQEAKEKDKMLALSLE 134
 QY 111 QESQRHGDIIQKDFLDVYVNLTKTMGIEWVHRFCPOAAFYVMTKDSMFINDVYLTLL 170
 DB 135 DEHLLYGDIIRQDFLDYNNLTLMKIMAFRWVTEFCPNKAYVMTKDTDFINTGNLVKYL 194
 QY 171 LKKNRTTFFTGFLKNEFPPIQPFSEKWFVSKSEYPMWDRYPPFCSGTGVYFGDVASQVY 230
 DB 195 LNLHSEKFTGYPLIDNYSYRGFYQKTHISYQEPFKVFPYPCSGLYIMSDLPRIY 254
 QY 231 NVSKSVPIKLEDVFGVGLCLERLNIRLELHQSPTFFPGGLRFSVCLFRRIVACHFIKPR 290
 DB 255 EMGHVVKPIKEDVYVVGICLNLKLVNIHIPEDTNLFLLYRIHLDVQCLRRVIAHGFSSK 314
 QY 291 TLLDYWQAL 299
 DB 315 EIITFWQVM 323

RESULT 11

AAU80224
 ID AAU80224 standard; Protein; 331 AA.

XX AAU80224;

XX 30-JUL-2002 (first entry)

XX Human betal,3-acetylglactosamine transferase.

XX Human; enzyme: betal,3-acetylglactosamine transferase;
 XX Gb4 sugar chain.

XX Homo sapiens.

XX JP2002085069-A.

XX 26-MAR-2002.

XX 08-SEP-2000; 2000JP-0273835.

XX 08-SEP-2000; 2000JP-0273835.

XX (SEKK) SEIKAGAKU KOGYO CO LTD.
 XX (FURU) FURUKAWA K.

XX WPI; 2002-378274/41.

XX N-PSDB; ABK51201.

XX Preparation of beta-1,3-acetylglactosamine transferase -
 PT ClaIm 3; Page 13-14; 15pp; Japanese.

XX The invention relates to the preparation of beta-1,3-acetylglactosamine
 CC (a Gb4 sugar chain molecule) comprising introducing a DNA appearing as
 CC ABK51201 (S1) or a DNA hybridising with S1 or a base sequence
 CC complementary to the base sequence or part of these base sequences under
 CC a stringent condition into a cell and growing the cell to express beta-1,
 CC 3-acetylglactosamine transferase and collecting it. Also included is the
 CC preparation of Gb4 sugar chain comprising contacting a polypeptide
 CC comprising the beta-1,3-acetylglactosamine transferase protein appearing
 CC as AAU80224 (A) or a polypeptide consisting of an amino acid sequence in
 CC which at least one amino acid is replaced, deleted, inserted or
 CC transferred in the amino acid sequence (A) and having enzymatic activity
 CC transferring N-acetylglactosamine residue from an N-acetylglactosamine
 CC donor to the C3 site of the galactose residue in Gb3 sugar chain which is
 CC the receptor with an N-acetylglactosamine donor and Gb3 sugar chain.
 CC The present sequence represents beta-1,3-acetylglactosamine transferase.

XX Sequence 331 AA;

Query Match 33.5%; Score 555.5; DB 23; Length 331;
 Best Local Similarity 36.2%; Pred. No. 7.4e-54;
 Matches 112; Conservative 57; Mismatches 125; Indels 15; Gaps 2;

QY 6 MRLMYICLLVLCALCLYFSMYSLNP-----FKEQSFVYKKGDNFLKLPDTCR 53
 DB 15 LRSKWSLLLSLSSFFVMYLSLPHYNVIERVNMWYFYEPIYRQDFHFTLREHSNCS 74
 QY 54 QTPPELVLLVSSHKQLAERMAIROTWGKRMVKGKOLKTFELLGTTSSAAE---TKEVD 110
 DB 75 HONPELVILVTHSPSDVKARQAIRVTWGEKSWGVEVLTFLLGQEAKEKDKMLALSLE 134
 QY 111 QESQRHGDIIQKDFLDVYVNLTKTMGIEWVHRFCPOAAFYVMTKDSMFINDVYLTLL 170
 DB 135 DEHLLYGDIIRQDFLDYNNLTLMKIMAFRWVTEFCPNKAYVMTKDTDFINTGNLVKYL 194
 QY 171 LKKNRTTFFTGFLKNEFPPIQPFSEKWFVSKSEYPMWDRYPPFCSGTGVYFGDVASQVY 230
 DB 195 LNLHSEKFTGYPLIDNYSYRGFYQKTHISYQEPFKVFPYPCSGLYIMSDLPRIY 254
 QY 231 NVSKSVPIKLEDVFGVGLCLERLNIRLELHQSPTFFPGGLRFSVCLFRRIVACHFIKPR 290
 DB 255 EMGHVVKPIKEDVYVVGICLNLKLVNIHIPEDTNLFLLYRIHLDVQCLRRVIAHGFSSK 314
 QY 291 TLLDYWQAL 299
 DB 315 EIITFWQVM 323

RESULT 12

AAU80224
 ID AAU80224 standard; Protein; 331 AA.

XX AAU80224;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:14093.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241859.
XX (HELI-) HELIX RES INST.
PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI: 2001-318749/34.
DR
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs.
XX
XX
PS Claim 8; SEQ ID 14093; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH9893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX
SQ Sequence 331 AA;

Query Match 33.2%; Score 550.5; DB 22; Length 331;
Best Local Similarity 36.2%; Pred. No. 2.7e-53;
Matches 112; Conservative 56; Mismatches 126; Indels 15; Gaps 2;
QY 6 MRLMYICLLVIGALCLYFSMYSNP-----FKQSFYVKKDGNFLKLPDTCR 53
DB 15 LRSLSKLSLLLSLFFVMYLSLPHYVIERVNMWTFYEPYRQDFHTLREHNSCS 74
QY 54 QTPPFLVLLVSSHKOLAERMAIQTWGMKRMVKGKOLKTFLLGTTSSAAE---TKREVD 110
DB 75 HQNPFLVILVTSHPSDVKARQAVRTWGEKKSWMGYEVLTFFLLGQAEKEDKMLALSLE 134
QY 111 QESQRGDIQKDFLDVYVNLTKTMGIEWHRRFCQPAAFVMTDSMDINVDYLTLL 170
DB 135 DEHLLYGDIIRQDFLOTNLTNLTNAFRWTFEFCNPAKVMKTDVDFINTGNLYKL 194
QY 171 LKNRRTRFTGTGLKLNKEFFIRQPFKSWFVSKSEYPWDYPPFCSGYGVFSGDVASQVY 230
DB 195 LNLNHEKFTFTPLDINYSYRGFYQKTHISYQEPYFVSPFYCSGLGYIMSRDLVPRIY 254
QY 231 NVSKSVPIYKLEVDVFGVGLCLERLNIRLEELHSQPTFFPGGLRFVSCLFRIVACHFIKPR 290
DB 255 EMAGHVKPIKFDVYVYICLNLKLVNIHIPEDNLNLPFLYRIHLDVCLQRRVIAAHGFSSK 314
QY 291 TLLDYQWAL 299
DB 315 EIITFWQVM 323

RESULT 13
AAE15927
ID AAE15927 standard; Protein; 193 AA.
XX
AC AAE15927;
XX
XX 26-MAR-2002 (first entry)
XX Human 4566767/84-303 protein, member of Fringe and Brainiac family.
XX
XX Fringe protein; Brainiac protein; glycosyltransferase; EGF; mannose;
KW N-acetylglucosamine; fucose; epidermal growth factor; T cell leukaemia;
KW breast cancer; stroke; dementia; leucoencephalopathy; Alagille syndrome;
KW notch protein; EGF-module containing protein; human.
XX
OS Homo sapiens.
XX
XX WO200187321-A2.
XX
XX 22-NOV-2001.
XX
XX 21-MAY-2001; 2001WO-IB01033.
XX
XX 19-MAY-2000; 2000GB-0012216.
XX
XX (EUMO-) EURO MOLECULAR BIOLOGY LAB.
XX
XX Cohen S, Brueckner K, Clausen H, Keck B;
XX WPI: 2002-082940/11.
XX
XX Use of Fringe protein or Brainiac protein as glycosyltransferase and
PT treatment of disease caused by epidermal growth factor-like module
PT containing protein, such as T cell leukemia, breast cancer, stroke and
PT dementia
XX
XX Claim 7; Fig 1; 52pp; English.
XX
XX The invention relates to the use of a Fringe protein or a Brainiac
CC protein or a fragment or functional equivalent of a Fringe or a
CC Brainiac protein, as glycosyltransferase. Fringe and Brainiac
CC proteins have been found to possess glycosyltransferase activity in
CC transferring sugar residues onto certain proteins of biological
CC interest, so affecting the binding of effector molecules to these
CC proteins. Particularly they are useful for transferring a
CC N-acetylglucosamine moiety onto a fucose or a mannose substrate.
CC whether free or attached to a lipid, carbohydrate or protein. They
CC act as glycosyltransferase on notch protein (EGF-module containing
CC protein) family. They are useful in treatment of a disease caused
CC by epidermal growth factor (EGF)-like module containing protein,
CC such as T cell leukaemia, breast cancer, stroke, dementia, cerebral
CC autosomal dominant arteriopathy with subcortical infarcts,
CC leucoencephalopathy and Alagille syndrome. Fringe and Brainiac
CC proteins are also useful for screening ligands capable of modulating
CC the activity of the protein. The present sequence is human protein
CC which is a member of Fringe and Brainiac protein family.
XX
SQ Sequence 193 AA;

Query Match 30.0%; Score 497.5; DB 23; Length 193;
Best Local Similarity 49.7%; Pred. No. 1.2e-47;
Matches 96; Conservative 34; Mismatches 60; Indels 3; Gaps 2;
QY 63 VTSKHKOLAERMAIQTWGMKRMVKGKOLKTFLLGTTSSAAEKEVDSESRHGDIQK 122
DB 1 ISTHKFEADQAQIRETWGDNENFKGIATFLGLGNADPVLNOMVQESQIFHDIVE 60
QY 123 DFLDYYVNLTKTMGIEWHRRFCQPAAFVMTDSMDINVDYLTLLKKNR--TTRFF 180
DB 61 DFIDSYHNLTKLTMGMRWVATFCSKAKYVMTDSDFVNDNLIYLLKPTSPRRRYF 120
QY 181 TGLFLKNEFFIRQPFKSWFVSKSEYPWDYPPFCSGYGVFSGDVASQVNVSKSVPIK 240
DB 181 TGLFLKNEFFIRQPFKSWFVSKSEYPWDYPPFCSGYGVFSGDVASQVNVSKSVPIK 240


```

Query Match      29.08; Score 480; DB 23; Length 196;
Best Local Similarity 48.18; Pred. No. 1.2e-45;
Matches 91; Conservative 36; Mismatches 58; Indels 4; Gaps 2;

QY 69 QLAERMAIROTWGKRVKQKLTFFLLGTT--SSAAETKEVDQESORHGDIIOKDFLD 126
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 7 QIARRAIRQWGNESLAPGIIQITRIFFLLGLSLKNGYLQRAILEESRQYHDIIOQEYLD 66

QY 127 VYNNLTLMGIEWHVFQAFVWKTSDMFINDVYLTTELLKKNRTR--FFTGF 184
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 67 TYNNLTITLGMNHWATYCPHIPYVWKTSDMEVNTTEYLINRLKPLPPRHNYFTGYL 126

QY 185 KLNFFPIRQPFKWFVSKSEVPMDRYPPFCGTCGYVFGSDVASOVYVYKSVPIKLEDV 244
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 127 MRGYAPNRNKDSKWMPDLYPSERYPFCGTCGYVFGSDLAEXIFXVSLGIRLHLEDV 186

QY 245 FVGLCLERL 253
   | | | | | | | |
Db 187 YVGICLAKL 195

```

Search completed: April 14, 2003, 12:10:27
 Job time : 77 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2003, 12:08:21 ; Search time 46 Seconds

(without alignments)
647.862 Million cell updates/sec

Title: US-09-914-152-1

Perfect score: 1657
Sequence: 1 MAFPKRLMYICLLVLGALC.....TLDDYWOALNSRGEDCPVP 310

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 295.5 | 17.8 | 684 | 2 F86394 | protein T24P13.20 |
| 2 | 267 | 16.1 | 739 | 2 G86397 | protein T7N9.18 [i |
| 3 | 265 | 16.0 | 739 | 2 T10648 | hypothetical prote |
| 4 | 243.5 | 14.7 | 642 | 2 D96777 | hypothetical prote |
| 5 | 233.5 | 14.1 | 262 | 2 T20203 | hypothetical prote |
| 6 | 223.5 | 13.5 | 384 | 2 T20446 | hypothetical prote |
| 7 | 219.5 | 13.2 | 507 | 2 T18639 | hypothetical prote |
| 8 | 209.5 | 12.6 | 345 | 2 T20031 | hypothetical prote |
| 9 | 200 | 12.1 | 357 | 2 T20029 | hypothetical prote |
| 10 | 199 | 12.0 | 376 | 2 T24925 | hypothetical prote |
| 11 | 192 | 11.6 | 332 | 2 T20028 | hypothetical prote |
| 12 | 189.5 | 11.4 | 376 | 2 T20876 | hypothetical prote |
| 13 | 171.5 | 10.4 | 398 | 2 B86353 | hypothetical prote |
| 14 | 160 | 9.7 | 353 | 2 C96573 | protein F2E2.6 [im |
| 15 | 158.5 | 9.6 | 253 | 2 T23387 | hypothetical prote |
| 16 | 155.5 | 9.4 | 325 | 2 T24762 | hypothetical prote |
| 17 | 146.5 | 8.8 | 409 | 2 A84733 | hypothetical prote |
| 18 | 143.5 | 8.7 | 404 | 2 A86186 | hypothetical prote |
| 19 | 141.5 | 8.5 | 359 | 2 F87937 | hypothetical prote |
| 20 | 141.5 | 8.5 | 385 | 2 T20879 | protein F14B6.6 [i |
| 21 | 140.5 | 8.5 | 338 | 2 T24743 | hypothetical prote |
| 22 | 139.5 | 8.4 | 399 | 2 H86453 | hypothetical prote |
| 23 | 139 | 8.4 | 406 | 2 T04817 | avr9 homolog F9L11 |
| 24 | 134 | 8.1 | 395 | 2 A86458 | hypothetical prote |
| 25 | 132.5 | 8.0 | 348 | 2 T31918 | probable elicitor |
| 26 | 126.5 | 7.6 | 334 | 2 T32256 | hypothetical prote |
| 27 | 122.5 | 7.4 | 401 | 2 A86251 | hypothetical prote |
| 28 | 109.5 | 6.6 | 333 | 2 T02614 | hypothetical prote |
| 29 | 105.5 | 6.4 | 378 | 2 T39790 | hypothetical prote |

| | | | | | |
|----|-------|-----|------|----------|--------------------|
| 30 | 105.5 | 6.4 | 753 | 2 T24745 | hypothetical prote |
| 31 | 105.5 | 6.4 | 1049 | 2 S67613 | vacuolar carboxype |
| 32 | 104 | 6.3 | 414 | 2 B96808 | protein F28K19.2 [|
| 33 | 101 | 6.1 | 349 | 2 T24744 | hypothetical prote |
| 34 | 100 | 6.0 | 845 | 2 T46164 | nodulin / glutamat |
| 35 | 98 | 5.9 | 660 | 2 T22794 | hypothetical prote |
| 36 | 95.5 | 5.8 | 347 | 2 D71690 | hypothetical prote |
| 37 | 94.5 | 5.7 | 3724 | 2 T18427 | hypothetical prote |
| 38 | 93.5 | 5.6 | 336 | 2 S70708 | site-specific DNA- |
| 39 | 93 | 5.6 | 368 | 2 T15096 | hypothetical prote |
| 40 | 92.5 | 5.6 | 279 | 2 T20443 | hypothetical prote |
| 41 | 90.5 | 5.5 | 338 | 2 T49998 | hypothetical prote |
| 42 | 90.5 | 5.5 | 1314 | 2 S19488 | probable membrane |
| 43 | 90 | 5.4 | 482 | 2 F82828 | phosphomannose iso |
| 44 | 90 | 5.4 | 786 | 2 T08664 | Toll protein-like |
| 45 | 88.5 | 5.3 | 579 | 2 S46757 | hypothetical prote |

ALIGNMENTS

RESULT 1

F86394 protein T24P13.20 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: F86394
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alo ansen, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzi Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tall ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F86394
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-684 <STO>
A:Cross-references: GB:AE005172; NID:g9295733; PIDN:AAF87039.1; GSPDB:GN00141
C:Genetics:
A:Gene: T24P13.20
A:Map position: 1

Query Match 17.8%; Score 295.5; DB 2; Length 684;
Best Local Similarity 26.3%; Pred. No. 1.4e-18;
Matches 92; Conservative 58; Mismatches 145; Indels 55; Gaps 10;

| | | | |
|----|-----|--|-----|
| QY | 9 | MYICLLVLGALCLYFSM-----YSLNPFKEQSFYKKDCNFKL-----LP----- | 48 |
| Db | 335 | LYIFLLLLAIFHQMFVSYGFFVQTFLEPLVSEIRITGDFRLISILASGLPTSESEHV | 394 |
| QY | 49 | -DTCRQTPTPF-----LVLLVTSKOLAERMAIQTWGKRMVKGKOLKTKFFLLGTT | 100 |
| Db | 395 | VDLEALKSPTSLRPLDLVIGVFSTANNFKRMVRRTWQYDDVRSGRVAVRFPVGLH | 454 |
| QY | 101 | SSAAETKEVDQSQRHGDIIQKDFLDVYVNLTKTM-----MGIEWVHRFC----- | 146 |
| Db | 455 | KSLPNLELWNEARTYGDVQLMPFVDYISLSWKTLAICIFGLSSLYCFASLTSYFKQTE | 514 |
| QY | 147 | -PQAAFMVKTDSDMFINDVLTPELLKKKRTTRFFTGFLKLNFPFIRQPSKWFVS---- | 201 |
| Db | 515 | VDSAFIMKTDGDAFVYRDEVLLSLSMNTNTRGLIYGLNSDSQPIRNPDSKWYISYEFL | 574 |
| QY | 202 | --KSEYPMWRDYPFCGSGTYGVFGSDVASOVYVNSK--SVPIKLEDVYFVGLCLERNLRL | 257 |
| Db | 575 | LKKQEMPEKYPFWAHGPGYIYVSRDIAESVGLFKEGNLMKFLEDVANGIWTAEUTKKG | 634 |
| QY | 258 | EELHSQPTFFPGGLRFSVCLFRRIYVACHFIKPRTLDDYWOALNSRGEDC | 307 |

Db 635 LEPHYEN---DGRISDCKDGYVVA-HYQSPAETCLWRKYQETKRSIC 680

RESULT 2

g86397

protein T7N9.18 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: G86397

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzila, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: G86397

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-657 <STO>

A:Cross-references: GB:AE005172; NID:98778858; PIDN:AAF79857.1; GSPDB:GN00141

C:Genetics:

A:Gene: T7N9.18

A:Map position: 1

Query Match 16.1%; Score 267; DB 2; Length 657;

Best Local Similarity 27.5%; Pred. No. 5e-16;

Matches 67; Conservative 51; Mismatches 102; Indels 24; Gaps 7;

Qy 73 RMAITQWAGRWKQKLTFFLLGTTSSAAETKEVDQSORHGDIIQKDFLDVYNYLT 132

Db 425 RMAVRKSMQOKLRSSKVAARFFVALHARKEVNVLDLKEAEYGDIVIVPMYDHYLVV 484

Qy 133 LKTM---MGTEWHRFPCQAFAAFVMTDSOMFINVDYLTLLKKNTTFFTCFLKINE 188

Db 485 LKTVACEYGVNTV----AAKYMKCDDDTFVRDVAIQAEKVGRESLYGINFNH 539

Qy 189 FPIQPSKFWVSKSEYPMWRYPFGSGTGVFGSDVASOVN--VSKSVPIKLEDFV 246

Db 540 KPLRT--GKVAVTFPEWPEEYPPYANGPGYILSYDVAKEFIVDDFEQKRLRFKMEVDVM 597

Qy 247 GLCLERINIRLEELHSOPTFFPGGLRFSVCLFRRI---VACHFTKPTLLDYWOALENSR 303

Db 598 GMWVEKFN-----ETRPVAVVHSLKF--COFGCIEDYFTHYQSPROMICMWDKLRQLG 649

Qy 304 GEDC 307

Db 650 KPQC 653

RESULT 3

T10648

hypothetical protein T13K14.220 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 15-Oct-1999

C:Accession: T10648

R:Bevan, M.; Pohl, T.; Weizenecker, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lemcke submitted to the Protein Sequence Database, June 1999

A:Reference number: Z16991

A:Accession: T10648

A:Molecule type: DNA

A:Residues: 1-739 <BEV>

A:Cross-references: EMBL:AL080282; GSPDB:GN00062; ATSP:T13K14.220

A:Experimental source: cultivar Columbia; BAC clone T13K14

C:Genetics:

A:Gene: ATSP:T13K14.220

A:Map position: 4

A:Introns: 59/3; 86/1; 358/1; 438/3; 535/3; 578/3; 634/3; 670/3

Query Match 16.0%; Score 265; DB 2; Length 739;

Best Local Similarity 26.8%; Pred. No. 8.8e-16;

Matches 79; Conservative 54; Mismatches 138; Indels 24; Gaps 9;

Qy 18 ALCLYFSMYSNLPFKEQSFVYKKGDNFLKLPDTCROTQPPF-LVLLVYSSHKQLAERMAI 76

Db 461 ATSLTSTSPSPSP--OKAIEFSSEWKAPPLPOT-----PFLFMGVLSATNHFSEMAV 512

Qy 77 ROTGKERWAGKQKLTFFLLGTTSSAAETKEVDQSORHGDIIQKDFLDVYNYLTLM 136

Db 513 RKTWQHPSIKSSDVVAFVFNPRKEVNAWLKAEYFGDIVILPMDRYELVVLTKI 572

Qy 137 MGIEVHRFCQAAPVMTKDSMFINDYLTLLKKNRTTFRFFFLKLEFPPIOPPS 196

Db 573 AICEFGNVTAP---YIMKCDDDTFRVESILKQIDGVSPKSLYMGNLNLRHRLPT--G 627

Qy 197 KWFYSKSEYPMWRYPFGSGTGVFGSDVASOV--YNYKSVPIKLEDFVGLCLERLN 254

Db 628 KWTVTWEEMPEAVPPYANGPGYIISNIAKYIVSONSRHKLRLFKMEDVSNGLMVEQFN 687

Qy 255 IRLEELHSOPTFFPGGLRFSV--CLFRIVACHFTKPTLLDYWOALENSRGDC 307

Db 688 ASM-----QPVEYSHSNKFCQYGCTLNYTA-HYQSPSQMCLMDNLLKGRPQC 736

RESULT 4

D96777

hypothetical protein F25A4.23 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: D96777

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzila, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: D96777

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-642 <STO>

A:Cross-references: GB:AE005173; NID:95882743; PIDN:AAD55296.1; GSPDB:GN00141

C:Genetics:

A:Gene: F25A4.23

A:Map position: 1

Query Match 14.7%; Score 243.5; DB 2; Length 642;

Best Local Similarity 24.9%; Pred. No. 6.5e-14;

Matches 65; Conservative 51; Mismatches 116; Indels 29; Gaps 8;

Qy 58 FLVLLVTSSHKQLAERMAIROTCKERVKGKOLKTFLLGTTSSAAETKEVDQESORHG 117

Db 396 FIGILSAGNH--FSEMAVRKSMQHVLTISAKVAVRFFVALHGRKEVNVLEKKEAYFG 453

Qy 118 DIIOKDFLDVYNYLTMTMGIEVHRFCQA-----FVAKTDSOMFINVDYLTLL 171

Db 454 DIVLPYMDSTDVVLVLTV-----AICEHGALAFSAKYIMKCDDDTFVVLGAVINEVK 506

Qy 172 KKNRTTFFTGFLKLEFPPIQPSKFWVSKSEYPMWRYPFGSGTGVFGSDVASOVYN 231

Db 507 KVPGRSLYIGNNYYHKPLRG--GKVAVTFPEWPEEYPPYANGPGVLUSSDIARFVD 564

Qy 232 VSK--SVPTYKLEDFVGLCLERINIRLEELHSOPTFFPGGLRFSVCLF---RRIVACHF 286

Db 565 KFERHKLRLFKMEDVSVGMVVEHF-----KNTTNPVDRHSLEF--COFGCVENYTAHY 617

Qy 287 IKPRTLLDYWOALENSRGDC 307

[illegible]

submitted to the EMBL Data Library, December 1997
A:Reference number: Z19212
A:Accession: T20031
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-345 <WIL>
A:Cross-references: EMBL:AL009246; PIDN:CAA15841.1; GSPDB:GN00019; CESP:C47F8.3
A:Experimental source: clone C47F8
R:Kershaw, J.
submitted to the EMBL Data Library, November 1996

A:Reference number: Z19617
A:Accession: T22806

A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA

A:Residues: 1-345 <WIL>
A:Cross-references: EMBL:Z81553; PIDN:CAB04502.1; GSPDB:GN00019; CESP:C47F8.3

A:Experimental source: clone F56H6
A:Genetics:
A:Gene: CESP:C47F8.3

A:Map position: 1
A:Introns: 23/1; 49/1; 169/3; 236/2; 278/3; 308/3

A:Superfamily: Caenorhabditis elegans hypothetical protein T09F5.1
Query Match 12.68; Score 209.5; DB 2; Length 345;
Best Local Similarity 24.68; Pred. No. 3.5e-11;
Matches 73; Conservative 46; Mismatches 129; Indels 49; Gaps 9;

Qy 11 ICLVLGALCLYFMSYSLNPFKEQSFYVK----- 39
Db 16 ICLLV-----FLFWKLPNPKSSFLEEHCVGSGNISTLSPRNDIFGSSFIVSPANI 69
Qy 40 -KCGNFKLPDTCRQTPPFLVLTSSHKQLAERMAIRQW---GKERMKVKQLKTFP 95
Db 70 HKSQKYLPRKMKVNTWKHDITLMLVSKTKNFARNVLRSTWKNKSENKMKSGRMHALF 129
Qy 96 LLGTTSSAAETKE-VDQESORHGDIIOKDFLDVYVNTLTMTMGLEWHRCPCQAAFYVK 154
Db 130 FVLGVPDGNLKLVLAEAEHGDVVDLEDYDNLPEFTLALLYGTSKASQFKIIGK 189
Qy 155 TDSDFINVDYLTLLKK--NRTTRFTGFLK-LNEFPIQPFKSWFVSKSEYPMWDRYP 211
Db 190 IDDDVMPFPDQLLPLDRNVNSNTLSYGHLSLTAELVLANKTEPMVVPYPTAYNCTVYP 249
Qy 212 PFCSGTGVVFGSDVASQVYVNVSKSVPIKLEDFV-GLCLERLNIRLEELHSQTFP 267
Db 250 VYVNGPIYLTDAASULDNANHQFMTVEDALITGIIAQKLGIR---RYSLPNVF 303

RESULT 9
T20029

hypothetical protein C47F8.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C:Accession: T20029

R:McLay, K.
submitted to the EMBL Data Library, December 1997

A:Reference number: Z19212
A:Accession: T20029

A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA

A:Residues: 1-357 <WIL>
A:Cross-references: EMBL:AL009246; PIDN:CAA15839.1; GSPDB:GN00019; CESP:C47F8.6

A:Experimental source: clone C47F8
A:Genetics:
A:Gene: CESP:C47F8.6

A:Map position: 1
A:Introns: 19/1; 52/1; 172/3; 239/2; 281/3; 311/3

A:Superfamily: Caenorhabditis elegans hypothetical protein T09F5.1
Query Match 12.18; Score 200; DB 2; Length 357;
Best Local Similarity 25.78; Pred. No. 2.7e-10;
Matches 59; Conservative 48; Mismatches 115; Indels 8; Gaps 4;

Qy 39 KKGNFLLKLPDTCRQTPPFLVLTSSHKQLAERMAIRQW---GKERMKVKQLKTFP 95
Db 73 QKNYTWLYPLTIEETSKETDILMIVASRTDSYARRNIMRQTMNKSNSSEIVANGRMKSUF 132
Qy 96 LLGTTSSAAETKE-VDQESORHGDIIOKDFLDVYVNTLTMTMGLEWHRCPCQAAFYVK 154
Db 133 LVGLAPADYKVKVMQBAKLYGDIIVDDMTDYBELIYKSLMIFLFGVSKAPQYKIIGK 192
Qy 155 TDSDFINVDYLTLLKK--NRTTRFTGFLK-LNEFPIQPFKSWFVSKSEYPMWDRYP 211
Db 193 IDEIMFFPDKLMALYEQGIIDSTPVSILYGLVIPAGRIDFRDKTRNRYVPEAYSCSQYP 252
Qy 212 PFCSGTGVVFGSDVASQVYVNVSKSVPIKLEDFV-GLCLERLNIRLEEL 260
Db 253 AYLSGMYMATREAAQMLLKSTKHDFIQVEDVLLTGILAEGLGPIRDM 302

RESULT 10
T24925

hypothetical protein T15D6.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
C:Accession: T24925

R:Dobson, R.
submitted to the EMBL Data Library, November 1996

A:Reference number: Z19956
A:Accession: T24925

A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA

A:Residues: 1-376 <WIL>
A:Cross-references: EMBL:Z83125; PIDN:CAB05616.1; GSPDB:GN00019; CESP:T15D6.5

A:Experimental source: clone T15D6
A:Genetics:
A:Gene: CESP:T15D6.5

A:Map position: 1
A:Introns: 59/1; 87/1; 207/3; 274/2; 316/3; 339/3

A:Superfamily: Caenorhabditis elegans hypothetical protein T09F5.1
Query Match 12.08; Score 199; DB 2; Length 376;
Best Local Similarity 23.58; Pred. No. 3.5e-10;
Matches 64; Conservative 60; Mismatches 120; Indels 28; Gaps 10;

Qy 39 KKGNFLLKLPDTCRQTPPFLVLTSSHKQLAERMAIRQW---GKERMKVKQLK 92
Db 108 QKSQRLHLPKI---KNPTNRDILMIVASRPGSVSRKRLKTKMKNKANSKIIRGRMQ 164
Qy 93 TFEELGTTSSAAE-TKEVDQESORHGDIIOKDFLDVYVNTLTMTMGLEWHRCPCQAA 151
Db 165 VFLVGVAGDRDLMAKAVKKEAESFGDIIVNLEDYDNLPEFTLALLYGTNKAADFPI 224
Qy 152 VMKTDSDMFINVLYLTLLKK--NRTTRFTGFLK-LNEFPIQPFKSWFVSKSEYPMW 208
Db 225 IGKIDDDVIFPPDRLTLLDENVIDSSSYIYLSQDDELVVVRNETPWPYPTAYNCT 284
Qy 209 RYPPFCSTGTGVFGSDVASQVYVNVSKSVPIKLEDFV-GLCLERLNIRLEELHSQTFP 267
Db 285 KYPVYALGFFYLTITNKAANLIVENSFRQFMTVEDALITGIIAEGGLI---QRHSLPMT 341
Qy 268 PGLRFESVCLFRFRIVACHFIK--PTLLDYQ 297
Db 342 DG-----KKILSWHMSKRSDRQFLDFYQ 364

RESULT 11
T20028

hypothetical protein C47F8.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C:Accession: T20028

R:McLay, K.
submitted to the EMBL Data Library, December 1997

A:Reference number: Z19212
A:Accession: T20028

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-332 <WIL>

A:Cross-references: EMBL:AL009246; PIDN:CAAL5838.1; GSPDB:GN00019; CESP:C47F8.5

A:Experimental source: clone C47F8

C:Genetics:

A:Gene: CESP:C47F8.5

A:Map position: 1

A:Introns: 26/1; 146/3; 213/2; 255/3; 285/3

C:Superfamily: Caenorhabditis elegans hypothetical protein T09F5.1

Query Match 11.6%; Score 192; DB 2; Length 332;
Best Local Similarity 25.7%; Pred. No. 1.3e-09;
Matches 59; Conservative 45; Mismatches 118; Indels 8; Gaps 4;

Qy 39 KKGDFLPLDPCROT-----PPF-----LVLLVTSKQLAERMAIRQTW---GKRMVKGKQLKTF 95

Db 47 QRTYKWLPLPETDILVRSPDLMLWASRTDSFARRNVLAKTWNKKNYSEIVRGRKALF 106

Qy 96 LIGTSSAAETKE-VDOESQRHGDIIQKDFLDVYNNLTUKTMGIEWVHRFCPOAAAFVK 154

Db 107 LVGMVSEYRVRRIVMEEAALYKGMVVIDLEDTYDLPFKSLSLLYAVSKAPEFKVIG 166

Qy 155 TSDMFINDVLTLELLK---NRTTRFTGFLKNEFFIRQPFKWFVSKSEYPWDVYP 211

Db 167 IDEVNFPPDKLIPLDGKVIDPDAAAFYQQLKEGEPIVKRKAHMYVPDYAINCTGYP 226

Qy 212 PFCSGTGYVFGSDVASQVNVSKSPYIKLED-VFVGLCLERLINIRLEEL 260

Db 227 AYVAGPFYLATKRAKLVLFKTFQNFMTVEDSLTGILANDLGIPIKNL 276

RESULT 12

T20876

hypothetical protein F14B6.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000

C:Accession: T20876

R:White, S.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19338

A:Accession: T20876

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-376 <WIL>

A:Cross-references: EMBL:Z81502; PIDN:CAB04107.1; GSPDB:GN00019; CESP:F14B6.4

A:Experimental source: clone F14B6

C:Genetics:

A:Gene: CESP:F14B6.4

A:Map position: 1

A:Introns: 31/1; 69/1; 189/3; 256/2; 298/3; 328/3

C:Superfamily: Caenorhabditis elegans hypothetical protein T09F5.1

Query Match 11.4%; Score 189.5; DB 2; Length 376;
Best Local Similarity 23.6%; Pred. No. 2.6e-09;
Matches 68; Conservative 57; Mismatches 138; Indels 25; Gaps 10;

Qy 39 KKGDFLPLDPCROT-----PPF-----LVLLVTSKQLAERMAIRQTW---GK 82

Db 76 KDFGSFTISFADIQKTHSWLYLPKFSMSSEILMIVASRTDSFARRNVLAKTWNPN 135

Qy 83 ERNVKGKQLKTFLLGTT--SSAAETKEVDQSRHGDIIQKDFLDVYNNLTUKTMGIE 140

Db 136 SEIIRKGRKALFLVGMGTGDDSRKRVMEEARIGYDMVVVDLKDYTEELPFKSLTLL 195

Qy 141 WYHRCPOAAFMKVDSDMFINDVLTLELLK---NRTTRFTGFL-KLNEFFIRQPF 197

Db 196 YGTSKASEFKLIGKIDEDIMFPDKLIPLLEONLDPSESIYGLFABGGVYRDKER 255

Qy 198 WFKSKSEYPMWRYPFCSGTGVYVFGSDVASQVNVSKSPYIKLEDVFVGLCLERLINIR 257

Db 256 WFPVDSYTCGMFPYPYTGGLFLVYQDAAKKILNATKRIPIEDIALINGILAN-DCRI 314

Qy 258 EELHSQTPFPFPGGLRFSVCLFRRIVACHFIK-PRTLIDY-WQALENSR 303

Db 315 PRTH-LPEIYGDHSLTDETRGVLAWHTGKNDQOFLDYFYKLSER 361

RESULT 13

B86353

protein F2E2.6 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: B86353

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Al-

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzi

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tall

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: B86353

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-398 <STO>

A:Cross-references: GB:AE005172; NID:g9280694; PIDN:AAF86563.1; GSPDB:GN00141

C:Genetics:

A:Gene: F2E2.6

A:Map position: 1

Query Match 10.4%; Score 171.5; DB 2; Length 398;

Best Local Similarity 26.0%; Pred. No. 1.2e-07;

Matches 59; Conservative 45; Mismatches 80; Indels 43; Gaps 11;

Qy 58 FLVLLVTSKQLAERMAIRQTW---GK--ERMVKGKQLKTFLLG--TTSSAAETREVD 110

Db 127 FVIGINTAFSSRRKRDLSRETWMPQGEKLEKEKIGVIVKMGHSITSPNSMLDEID 186

Qy 111 QESQRHGDIIQKDFLDVYNNLTUKTMG1-----EWVHRFCPOAAFMKVDSDMFINDV 165

Db 187 SEDQYNDFFLDHVEGYNNLSAKTSFFSSAVAKW-----DAEFYVKIDDDVHVNLGT 240

Qy 166 LTELKLNKTRTFTGFLKNEFFI-----RQP-FSKWFSVSKSEYPMWRYPFCSG 216

Db 241 LASTLASHRSKPRVYICMKSQ--PVLTKTKTAKYREPEFWKFGEGNKY-----FRHA 291

Qy 217 TG--YVFGSDVASQVNVSKSPYIKLEDV-----FVGLCLERLINIR 256

Db 292 TGOIYAISKDLATYISNNQPIHLKYANEDVTLAGSWFGLVEVEQIDDR 338

RESULT 14

C96573

protein F12M16.19 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: C96573

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Al-

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzi

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tall

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: C96573

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-353 <STO>

Search completed: April 14, 2003, 12:11:20
Job time : 48 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2003, 14:41:40 ; Search time 33 Seconds
(Without alignments)
1748.279 Million cell updates/sec

Title: US-09-914-152-1_COPY_31_310
Perfect score: 1498
Sequence: 1 FKEQSFVYKKDGNFLKLPDT.....TLDDYQWALENSRGEDCPPV 280

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: sp.archaea:*
 - 2: sp.bacteria:*
 - 3: sp.fungi:*
 - 4: sp.human:*
 - 5: sp.invertebrate:*
 - 6: sp.mammal:*
 - 7: sp.mhc:*
 - 8: sp.organelle:*
 - 9: sp.phage:*
 - 10: sp.plant:*
 - 11: sp.rodent:*
 - 12: sp.virus:*
 - 13: sp.vertibrate:*
 - 14: sp.unclassified:*
 - 15: sp.rvirus:*
 - 16: sp.bacteriap:*
 - 17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-----------|---------------------|
| 1 | 1389.5 | 92.8 | 302 | 6 Q9N292 | Q9n292 pongo pygma |
| 2 | 577 | 38.5 | 326 | 4 Q9Y526 | Q9y526 homo sapien |
| 3 | 577 | 38.5 | 326 | 6 Q9MYM7 | Q9mym7 pongo pygma |
| 4 | 577 | 38.5 | 326 | 11 Q54904 | Q54904 mus musculus |
| 5 | 572 | 38.2 | 305 | 11 Q920V5 | Q920v5 mus musculus |
| 6 | 572 | 38.2 | 305 | 11 Q91V52 | Q91v52 mus musculus |
| 7 | 569 | 38.0 | 422 | 4 O43825 | O43825 homo sapien |
| 8 | 567 | 37.9 | 409 | 11 Q920V4 | Q920v4 mus musculus |
| 9 | 567 | 37.9 | 409 | 11 Q920V3 | Q920v3 mus musculus |
| 10 | 567 | 37.9 | 409 | 11 Q920V2 | Q920v2 mus musculus |
| 11 | 567 | 37.9 | 409 | 11 Q91VE9 | Q91ve9 mus musculus |
| 12 | 567 | 37.9 | 409 | 11 Q91V58 | Q91v58 mus musculus |
| 13 | 567 | 37.9 | 409 | 11 Q91V19 | Q91v19 mus musculus |
| 14 | 567 | 37.9 | 422 | 11 O54905 | O54905 mus musculus |
| 15 | 565 | 37.7 | 422 | 4 Q9BZQ9 | Q9bzq9 homo sapien |
| 16 | 555.5 | 37.1 | 331 | 11 Q920V1 | Q920v1 mus musculus |

| | | | | | |
|----|-------|------|-----|-----------|---------------------|
| 17 | 555.5 | 37.1 | 331 | 11 Q920V0 | Q920v0 mus musculus |
| 18 | 555.5 | 37.1 | 331 | 11 Q91V72 | Q91v72 mus musculus |
| 19 | 555.5 | 37.1 | 331 | 11 O54906 | O54906 mus spicile |
| 20 | 548.5 | 36.6 | 331 | 4 O75752 | O75752 homo sapien |
| 21 | 529.5 | 35.3 | 255 | 11 Q9CTE5 | Q9cte5 mus musculus |
| 22 | 433.5 | 28.9 | 412 | 13 Q8UWM2 | Q8uwm2 brachydanio |
| 23 | 429 | 28.6 | 204 | 4 Q8TDV1 | Q8tdv1 homo sapien |
| 24 | 413 | 27.6 | 372 | 4 Q9C0J2 | Q9c0j2 homo sapien |
| 25 | 406 | 27.1 | 390 | 13 Q8UWM1 | Q8uwm1 brachydanio |
| 26 | 402.5 | 26.9 | 372 | 11 Q9D722 | Q9d722 mus musculus |
| 27 | 401.5 | 26.8 | 372 | 11 Q8R0U2 | Q8r0u2 mus musculus |
| 28 | 396 | 26.4 | 378 | 4 Q9BYG0 | Q9byg0 homo sapien |
| 29 | 393 | 26.2 | 371 | 11 Q88178 | Q88178 rattus norv |
| 30 | 388 | 25.9 | 370 | 11 Q920U7 | Q920u7 mus musculus |
| 31 | 388 | 25.9 | 370 | 11 Q91VC1 | Q91vc1 mus musculus |
| 32 | 388 | 25.9 | 371 | 11 Q920F0 | Q920f0 mus musculus |
| 33 | 388 | 25.9 | 406 | 13 Q8UWM0 | Q8uwm0 brachydanio |
| 34 | 387.5 | 25.9 | 378 | 4 Q96024 | Q96024 homo sapien |
| 35 | 384 | 25.6 | 418 | 13 Q8UWM4 | Q8uwm4 brachydanio |
| 36 | 383 | 25.6 | 370 | 11 Q920U8 | Q920u8 mus musculus |
| 37 | 382.5 | 25.5 | 277 | 4 Q8TA24 | Q8ta24 homo sapien |
| 38 | 382.5 | 25.5 | 384 | 4 Q8TDX1 | Q8tdx1 homo sapien |
| 39 | 381.5 | 25.5 | 574 | 5 Q9VLS8 | Q9vls8 drosophila |
| 40 | 381.5 | 25.5 | 585 | 5 Q95TQ8 | Q95tq8 drosophila |
| 41 | 380 | 25.4 | 370 | 11 Q920U9 | Q920u9 mus musculus |
| 42 | 364 | 24.3 | 412 | 13 Q8UWM3 | Q8uwm3 brachydanio |
| 43 | 362 | 24.2 | 399 | 11 Q8V116 | Q8v116 mus musculus |
| 44 | 359.5 | 24.0 | 377 | 4 Q96EK0 | Q96ek0 homo sapien |
| 45 | 359 | 24.0 | 397 | 11 Q91V18 | Q91v18 mus musculus |

ALIGNMENTS

RESULT 1

Q9N292 ID Q9N292 PRELIMINARY; PRT; 302 AA.
AC Q9N292; 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE UDP-Gal:GlcNAc betal,3-galactosyltransferase 5 (Fragment).
GN BETAL_3-CALT_5.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ORAN-PO17;
RA Liu Y., Saitou N.;
RT "Silver Project."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB041417; BAA94502.1; .
DR InterPro; IPR002659; Galactosyl_T.
DR Pfam; PF01762; Galactosyl_T; 1.
KW Glycosyltransferase; transferase.
FT NON_TER 302 302
SQ SEQUENCE 302 AA; 35500 MW; 51A555C864465A90 CRC64;

Query Match 92.8%; Score 1389.5; DB 6; Length 302;
Best Local Similarity 96.7%; Pred. No. 7.1e-126;
Matches 263; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

| | | | |
|----|-----|---|-----|
| QY | 1 | FKEQSFVYKK-DGNFLKLPDTCRQTPFVLVLTSSHKQLAERMAIRQTGKERMVKGK | 59 |
| Db | 31 | FKEQSFVYKKEDGNFLKLPDTCRQTPFVLVLTSSHKQLAERMAIRQTGKERTYKGK | 90 |
| QY | 60 | QLKTFLLGTTSSAAETKEVDQESQRHGDIITQKDFLOWYYNLTKTMGIEWHVFQQA | 119 |
| Db | 91 | QLKTFLLGTTSSAAENKEVDQESQRHGDIITQKDFLDVYYNLTKTMGMEMWHVFQQA | 150 |
| QY | 120 | AFVMTKDSDFINVDYLTLELLKKNRTRRFTGTGFLKLNFFIRQPFKSWFVSKSEYPWDR | 179 |


```
Db 185 DNLVYKLLKPKSTKRRRYFTGYV-INGGPIRDRVRSKWYMPDRDLYPDSNYPFCSCGTGYIF 243
Qy 192 SGDVASOVYVNSKSPYIKLEDFVGLCLERLNIRLEELHSQTPFFPGG-----LRF5VC 246
Db 244 SADVAELIYKTSLSHTRLLHLEDVYVGLCLRLGLIH-----PFQNSGFNHMKWAYSILC 295
Qy 247 LFRIVACHFIKPRPTLLDYWQALNSRGDC 277
Db 296 RYRRVITVHQISPEEMHRIWDMSSKKHLRC 326

RESULT 4
054904 PRELIMINARY; PRT; 326 AA.
ID O54904
AC O54904
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE UDP-galactose:BETA-N-acetyl-glucosamine-BETA-1,3-galactosyltransferase
DE II (EC 2.4.1.-) (UDP-GAL:BETAGLGNAC BETA 1,3-GALACTOSYLTRANSFERASE-I)
DE (BETA3GALT-1).
GN B3GT1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC STRAIN=129/SVJ;
RX MEDLINE=98079027; PubMed=9417047;
RA Hennes T., Dinter A., Kuhnert P., Mattu T.S., Rudd P.M., Berger E.G.;
RT "Genomic cloning and expression of three murine UDP-galactose: beta-N-
acetylglucosamine beta1,3-galactosyltransferase genes.";
RL J. Biol. Chem. 273:58-65(1998).
CC -1- FUNCTION: THIS PROTEIN IS RESPONSIBLE FOR THE SYNTHESIS OF
CC COMPLEX-TYPE N-LINKED OLIGOSACCHARIDES IN MANY GLYCOPROTEINS A AS
CC WELL AS THE CARBOHYDRATE MOETIES OF GLYCOLIPIDS. GALT-1, GALT-II
CC AND GALT-III ENZYMES ACT ON THE SAME ACCEPTOR SUBSTRATES BUT WITH
CC DIFFERENT AFFINITY.
CC -1- CATALYTIC ACTIVITY: UDP-GALACTOSE + N-ACETYL-BETA-D-GLUCOSAMINYL-
CC GLYCOPOLYMERIDE -> UDP + BETA-D-GALACTOSYL-1,3-N-ACETYL-BETA-D-
CC GLUCOSAMINYLGLYCOPOLYMERIDE.
CC -1- COFACTOR: MANGANESE.
CC -1- PATHWAY: GLYCOSYLATION.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND TO A LESSER EXTENT IN
CC HEART, KIDNEY, LIVER, LUNG, OVARY, TESTIS, UTERUS, COLON, STOMACH,
CC BONE MARROW, THYMUS, SPLEEN AND LYMPH NODES.
CC EMBL: AF029790; AAC53523.1;
CC MGD: MGI:1349403; B3galt1.
CC InterPro: IPR002659; Galactosyl_T.
CC Pfam: PF01762; Galactosyl_T; 1.
KW Transferase; Glycosyltransferase; Transmembrane; Signal-anchor;
KW Glycoprotein; Golgi stack.
FT DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 7 25 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DOMAIN 26 326 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 47 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 151 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 326 AA; 37979 MW; 9825D68869EFAD3D CRC64;

Query Match 38.5%; Score 577; DB 11; Length 326;
Best Local Similarity 42.8%; Pred. No. 1.9e-47;
Matches 116; Conservative 46; Mismatches 93; Indels 16; Gaps 4;

Qy 14 FLKLPDTCRQTPPFLVLLTSSHKOLAERMAIROTWGKRMVKGKQLKTFLLGTSSA 73
Db 65 FLINPNKCEKNIPFLVILSTTHKEFDARQAIRETWGDENNFKGIKTIATLFLGKNADP 124
Qy 74 AETKEVDQESQRHGDIQKDFLDVYVNLTKTMGIEWHREFCQAAYVVKMTSDMFINV 133
Db 125 VLNQWVEQSQIFHDIIVEDFIDSYHNLTKLTLMGMRVATFCSKAKYVVKMTSDIFVNM 184
```

```
Qy 134 DYLTLLKLNKR--TTREFTGFLKNEPPIROFFSKWFKSEYWDRIYPPFCSCGTGYVF 191
Db 185 DNLVYKLLKPKSTKRRRYFTGYV-INGGPIRDRVRSKWYMPDRDLYPDSNYPFCSCGTGYIF 243
Qy 192 SGDVASOVYVNSKSPYIKLEDFVGLCLERLNIRLEELHSQTPFFPGG-----LRF5VC 246
Db 244 SADVAELIYKTSLSHTRLLHLEDVYVGLCLRLGLIH-----PFQNSGFNHMKWAYSILC 295
Qy 247 LFRIVACHFIKPRPTLLDYWQALNSRGDC 277
Db 296 RYRRVITVHQISPEEMHRIWDMSSKKHLRC 326

RESULT 5
0920V5 PRELIMINARY; PRT; 305 AA.
ID 0920V5
AC 0920V5
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase I (Fragment).
GN B3GT1.
OS Mus spicilegus (Steppe mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZBN;
RA Liu Y., Kitano T., Koide T., Shiroishi T., Moriaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
Five Mus musculus subspecies.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC EMBL: AB039143; BAB68667.1;
CC InterPro: IPR002659; Galactosyl_T.
CC Pfam: PF01762; Galactosyl_T; 1.
KW Glycosyltransferase; Transferase.
FT NON_TER 1
FT NON_TER 305
SQ SEQUENCE 305 AA; 35611 MW; A338019F3B06EF5A CRC64;

Query Match 38.2%; Score 572; DB 11; Length 305;
Best Local Similarity 44.2%; Pred. No. 5.4e-47;
Matches 115; Conservative 43; Mismatches 86; Indels 16; Gaps 4;

Qy 14 FLKLPDTCRQTPPFLVLLTSSHKOLAERMAIROTWGKRMVKGKQLKTFLLGTSSA 73
Db 54 FLINPNKCEKNIPFLVILSTTHKEFDARQAIRETWGDENNFKGIKTIATLFLGKNADP 113
Qy 74 AETKEVDQESQRHGDIQKDFLDVYVNLTKTMGIEWHREFCQAAYVVKMTSDMFINV 133
Db 114 VLNQWVEQSQIFHDIIVEDFIDSYHNLTKLTLMGMRVATFCSKAKYVVKMTSDIFVNM 173
Qy 134 DYLTLLKLNKR--TTREFTGFLKNEPPIROFFSKWFKSEYWDRIYPPFCSCGTGYVF 191
Db 174 DNLVYKLLKPKSTKRRRYFTGYV-INGGPIRDRVRSKWYMPDRDLYPDSNYPFCSCGTGYIF 232
Qy 192 SGDVASOVYVNSKSPYIKLEDFVGLCLERLNIRLEELHSQTPFFPGG-----LRF5VC 246
Db 233 SADVAELIYKTSLSHTRLLHLEDVYVGLCLRLGLIH-----PFQNSGFNHMKWAYSILC 284
Qy 247 LFRIVACHFIKPRPTLLDYW 266
Db 285 RYRRVITVHQISPEEMHRIW 304

RESULT 6
091V52 PRELIMINARY; PRT; 305 AA.
ID 091V52
AC 091V52
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
```

DE UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase I (Fragment).
GN B3Grl.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=1090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VARIOUS STRAINS;
RA Liu Y., Kitano T., Koide T., Shirolshi T., Moriaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
RT Five Mus musculus subspecies";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB039134; BAB68658.1; -
DR EMBL; AB039135; BAB68659.1; -
DR EMBL; AB039136; BAB68660.1; -
DR EMBL; AB039137; BAB68661.1; -
DR EMBL; AB039138; BAB68662.1; -
DR EMBL; AB039139; BAB68663.1; -
DR EMBL; AB039140; BAB68664.1; -
DR EMBL; AB039141; BAB68665.1; -
DR EMBL; AB039142; BAB68666.1; -
DR InterPro: IPR002659; Galactosyl_T.
DR Pfam: PF01762; Galactosyl_T.
KW Glycosyltransferase; Transferase.
FT NON_TER 1
FT NON_TER 305
SQ SEQUENCE 305 AA; 35611 MW; A338019F3806EF5A CRC64;

Query Match 38.2%; Score 572; DB 11; Length 305;
Best Local Similarity 44.2%; Pred. No. 5.4e-47;
Matches 115; Conservative 43; Mismatches 86; Indels 16; Gaps 4;

QY 14 FLKLPDTCRQTPFPFLVLLVTSKQLAERMAIRQTGKRMVKGKOLKTFLLGTSSA 73
DB 54 FLINENKCEKNIPFLVLLVTSKFEQDARQALRETWGDNFKGKIATFLGKNADP 113
QY 74 AETKEVDQSQRGDIQKDFLDVYNNLTGKMMGLWVHRCFQAAFYVKTDSDMFINV 133
DB 114 VLNQMVQEQSQIFHDIIEDFIDSYHNLTKLTMGRVATFCGKAKYVKTDSDFVNM 173
QY 134 DYLTELLAKNR--TTREFTGLKNEFFIROPFSKWFVSKSEYPDWRYPFCSGTGYVF 191
DB 174 DNLVLLKAPSKYPRRYFTGV--INGGPIRVRSKWMPDRDLYPDSNYPFCSGTGYIF 232
QY 192 SGDVASQVNVSKSVPIKLEDFVGLCLERLNIRLEELHSDPTPPGG-----LRFVYC 246
DB 233 SADVAELIVKTSILHLLHLEDDVYVGLCLRLGIIH-----PFQNSGFNHWKMAYSLC 284
QY 247 LFRIRVACHFIKPRLLDYW 266
DB 285 RYRRVITVHOISPEEMHRIW 304

RESULT 7
Q43825 PRELIMINARY; PRT: 422 AA.
ID O43825
AC O43825;
DC 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE UDP-galactose:BETA-N-acetyl-glucosamine-BETA-1,3-galactosyltransferase
DE 2 (EC 2.4.1.-) (BETA3GALT2).
GN GALT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC TISSUE-BRAIN;
RX MEDLINE=98079080; PubMed=9417100;
RA Kolbinger F., Streiff M.B., Katopodis A.G.;

RT *Cloning of a human UDP-galactose:2-acetamido-2-deoxy-D-glucose beta-
RT galactosyltransferase catalyzing the formation of type 1 chains.*;
RL J. Biol. Chem. 273:433-440(1998).
RN [2]
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC TISSUE-FETAL BRAIN;
RX MEDLINE=98250717; PubMed=9582303;
RA Amado M., Almeida R., Carneiro F., Levery S.B., Holmes E.H.,
RA Bennett E.P., Clausen H.;
RT A family of human beta3-galactosyltransferases. Characterization of
RT four members of a UDP-galactose:beta-N-acetyl-glucosamine/Beta-
RT nacyl-galactosamine beta-1,3-galactosyltransferase family.*;
RL J. Biol. Chem. 273:12770-12778(1998).
CC -!- FUNCTION: THIS PROTEIN IS RESPONSIBLE FOR THE SYNTHESIS OF
CC COMPLEX-TYPE N-LINKED OLIGOSACCHARIDES IN MANY GLYCOPROTEINS A AS
CC WELL AS THE CARBOHYDRATE MOETIES OF GLYCOLIPIDS. GALT1 AND GALT2
CC ENZYMES ACT ON THE SAME ACCEPTOR SUBSTRATES BUT WITH DIFFERENT
CC AFFINITY.
CC -!- CATALYTIC ACTIVITY: UDP-GALACTOSE + N-ACETIL-BETA-D-GLUCOSAMINYL-
CC GLYCOPETIDE = UDP + BETA-D-GALACTOSYL-1,3-N-ACETYL-BETA-D-
CC GLUCOSAMINYLGLYCOPETIDE.
CC -!- PATHWAY: GLYCOSYLATION.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN HEART AND TO A LESSER EXTENT IN
CC BRAIN. NO EXPRESSION IN PLACENTA, KIDNEY, LIVER, SKELETAL MUSCLE,
CC LUNG AND PANCREAS.
DR EMBL; Y15014; CAA75245.1; -
DR EMBL; Y15060; CAA75344.1; -
DR MIM: 603018; -
DR InterPro: IPR002659; Galactosyl_T.
DR Pfam: PF01762; Galactosyl_T.
KW Transferase; Glycosyltransferase; Transmembrane; Signal-anchor;
KW Glycoprotein; Golgi stack.
FT DOMAIN 1 24 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 25 43 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DOMAIN 44 422 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 75 75 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 100 100 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 176 176 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 226 226 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 422 AA; 49213 MW; BA62942D6AFB4B0C CRC64;

Query Match 38.0%; Score 569; DB 4; Length 422;
Best Local Similarity 41.5%; Pred. No. 1.6e-46;
Matches 108; Conservative 53; Mismatches 95; Indels 4; Gaps 2;

QY 22 CROTTPFLVLLVTSKQLAERMAIRQTGKRMVKGKOLKTFLLGT--SSAAETKEV 79
DB 146 COEKSPFLILLIAAEPQIEARRAIRQTGWSLAPGCIOTRIFLLGLSLKNCYLORAI 205
QY 80 DQESQRHGDIIQKDFLDVYNNLTGKMMGLWVHRCFQAAFYVKTDSDMFINVDTL 139
DB 206 LEESROYHDIQOEYLDYNNLTGKMMGLWVHRCFQAAFYVKTDSDMFINVDTL 265
QY 140 LLKKNRTTR--PFTGFLKNEFFIROPFSKWFVSKSEYPDWRYPFCSGTGYVFGDVAS 197
DB 266 LLKPDLPFRHNTFTGLMRGYAPNRKDKSWYMPDLYPSERYVPFCSGTGYVFGDLAE 325
QY 198 QYVNVSKSVPIKLEDFVGLCLERLNIRLEELHSDPTPPGGGLRFSVCLFRIRVACHFI 257
DB 326 KIFKVSGLIRRLHLEDVYVGLCLERLNIRLEELHSDPTPPGGGLRFSVCLFRIRVACHFI 395
QY 258 KPRTLDDYQWALENSRGEDC 277
DB 386 QPSELIKYWNHLQONKNAC 405

RESULT 8
Q920V4 PRELIMINARY; PRT: 409 AA.
ID Q920V4
AC Q920V4;

```
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE UDP-Gal:betaglcNac beta 1,3-galactosyltransferase II (Fragment).
GN B3GT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CAST/EI;
RA Liu Y., Kitano T., Koide T., Shiroishi T., Moriaki K., Saitou N.;
RT *Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
RL Five Mus musculus subspecies.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB039147; BAB68671.1; -
DR InterPro; IPR002659; Galactosyl_T.
DR Pfam; PF01762; Galactosyl_T; 1.
KW Glycosyltransferase; Transferase.
FT NON_TER 1
FT NON_TER 409
SQ SEQUENCE 409 AA; 47247 MW; 2425654F23A8F0E3 CRC64;

Query Match 37.9%; Score 567; DB 11; Length 409;
Best Local Similarity 41.2%; Pred. No. 2.4e-46;
Matches 107; Conservative 53; Mismatches 96; Indels 4; Gaps 2;

Qy 22 CROTPPFLVLLVTSKOLAERMAIRQTWGNETLAPGQIIRVFLGLGSIKUNGYLQHA 199
Db 140 COEKSPFLLILIAEPGQIEARRAIRTWGNETLAPGQIIRVFLGLGSIKUNGYLQHA 199
Qy 80 DOESORHGDIIOKQFLDYVYNTLTMTMGIEVHRCFQAAAFVMTKDSDFINVDYLT 139
Db 200 QEESQYHDIIOQEYLDYVYNTLTMTMGNNVATYCPHTPVYMTKDSDFINVDYLT 259
Qy 140 LLKNRTR--FFTGFLKNEFPPIQPFKSFVSKSEYPMWDRYPPFCSTGYVFGDVA 197
Db 260 LLKPDLPPIRHHNYFTGYLMRGYAPNRNKSWMYPPDLPYSERYPVFCSTGYVFGDLAE 319
Qy 198 QVNVSKVPYIKLEDVFGCLERINIRLEELHSQPTFFPGGLRFSVCLFRIVACHFI 257
Db 320 KIFKVSGLIRLHLEDVYVIGCLAKLRVDPVPPNFEVFNHVRVSYSSCKYSHLITSHQF 379
Qy 258 KPRTLLDYWQALENSRGDC 277
Db 380 QPSELIKYNNHLOQKNHAC 399

RESULT 9
Q920V3 ID Q920V3 PRELIMINARY; PRT; 409 AA.
AC Q920V3
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE UDP-Gal:betaglcNac beta 1,3-galactosyltransferase II (Fragment).
GN B3GT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZBN;
RA Liu Y., Kitano T., Koide T., Shiroishi T., Moriaki K., Saitou N.;
RT *Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
RL Five Mus musculus subspecies.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB039153; BAB68677.1; -
DR InterPro; IPR002659; Galactosyl_T.
DR Pfam; PF01762; Galactosyl_T; 1.
KW Glycosyltransferase; Transferase.
FT NON_TER 1
FT NON_TER 409
SQ SEQUENCE 409 AA; 47383 MW; 0DC51AB6335A58DE CRC64;

Query Match 37.9%; Score 567; DB 11; Length 409;
Best Local Similarity 41.2%; Pred. No. 2.4e-46;
Matches 107; Conservative 53; Mismatches 96; Indels 4; Gaps 2;

Qy 22 CROTPPFLVLLVTSKOLAERMAIRQTWGNETLAPGQIIRVFLGLGSIKUNGYLQHA 199
Db 140 COEKSPFLLILIAEPGQIEARRAIRTWGNETLAPGQIIRVFLGLGSIKUNGYLQHA 199
Qy 80 DOESORHGDIIOKQFLDYVYNTLTMTMGIEVHRCFQAAAFVMTKDSDFINVDYLT 139
Db 200 QEESQYHDIIOQEYLDYVYNTLTMTMGNNVATYCPHTPVYMTKDSDFINVDYLT 259
Qy 140 LLKNRTR--FFTGFLKNEFPPIQPFKSFVSKSEYPMWDRYPPFCSTGYVFGDVA 197
Db 260 LLKPDLPPIRHHNYFTGYLMRGYAPNRNKSWMYPPDLPYSERYPVFCSTGYVFGDLAE 319
Qy 198 QVNVSKVPYIKLEDVFGCLERINIRLEELHSQPTFFPGGLRFSVCLFRIVACHFI 257
Db 320 KIFKVSGLIRLHLEDVYVIGCLAKLRVDPVPPNFEVFNHVRVSYSSCKYSHLITSHQF 379
Qy 258 KPRTLLDYWQALENSRGDC 277
Db 380 QPSELIKYNNHLOQKNHAC 399
```

```
FT NON_TER 1
FT NON_TER 409
SQ SEQUENCE 409 AA; 47220 MW; 2D2EBC4D68E8463D CRC64;

Query Match 37.9%; Score 567; DB 11; Length 409;
Best Local Similarity 41.2%; Pred. No. 2.4e-46;
Matches 107; Conservative 53; Mismatches 96; Indels 4; Gaps 2;

Qy 22 CROTPPFLVLLVTSKOLAERMAIRQTWGNETLAPGQIIRVFLGLGSIKUNGYLQHA 199
Db 140 COEKSPFLLILIAEPGQIEARRAIRTWGNETLAPGQIIRVFLGLGSIKUNGYLQHA 199
Qy 80 DOESORHGDIIOKQFLDYVYNTLTMTMGIEVHRCFQAAAFVMTKDSDFINVDYLT 139
Db 200 QEESQYHDIIOQEYLDYVYNTLTMTMGNNVATYCPHTPVYMTKDSDFINVDYLT 259
Qy 140 LLKNRTR--FFTGFLKNEFPPIQPFKSFVSKSEYPMWDRYPPFCSTGYVFGDVA 197
Db 260 LLKPDLPPIRHHNYFTGYLMRGYAPNRNKSWMYPPDLPYSERYPVFCSTGYVFGDLAE 319
Qy 198 QVNVSKVPYIKLEDVFGCLERINIRLEELHSQPTFFPGGLRFSVCLFRIVACHFI 257
Db 320 KIFKVSGLIRLHLEDVYVIGCLAKLRVDPVPPNFEVFNHVRVSYSSCKYSHLITSHQF 379
Qy 258 KPRTLLDYWQALENSRGDC 277
Db 380 QPSELIKYNNHLOQKNHAC 399

RESULT 10
Q920V2 ID Q920V2 PRELIMINARY; PRT; 409 AA.
AC Q920V2
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE UDP-Gal:betaglcNac beta 1,3-galactosyltransferase II (Fragment).
GN B3GT2.
OS Mus spicilegus (Steppe mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZBN;
RA Liu Y., Kitano T., Koide T., Shiroishi T., Moriaki K., Saitou N.;
RT *Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
RL Five Mus musculus subspecies.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB039153; BAB68677.1; -
DR InterPro; IPR002659; Galactosyl_T.
DR Pfam; PF01762; Galactosyl_T; 1.
KW Glycosyltransferase; Transferase.
FT NON_TER 1
FT NON_TER 409
SQ SEQUENCE 409 AA; 47383 MW; 0DC51AB6335A58DE CRC64;

Query Match 37.9%; Score 567; DB 11; Length 409;
Best Local Similarity 41.2%; Pred. No. 2.4e-46;
Matches 107; Conservative 53; Mismatches 96; Indels 4; Gaps 2;

Qy 22 CROTPPFLVLLVTSKOLAERMAIRQTWGNETLAPGQIIRVFLGLGSIKUNGYLQHA 199
Db 140 COEKSPFLLILIAEPGQIEARRAIRTWGNETLAPGQIIRVFLGLGSIKUNGYLQHA 199
Qy 80 DOESORHGDIIOKQFLDYVYNTLTMTMGIEVHRCFQAAAFVMTKDSDFINVDYLT 139
Db 200 QEESQYHDIIOQEYLDYVYNTLTMTMGNNVATYCPHTPVYMTKDSDFINVDYLT 259
Qy 140 LLKNRTR--FFTGFLKNEFPPIQPFKSFVSKSEYPMWDRYPPFCSTGYVFGDVA 197
Db 260 LLKPDLPPIRHHNYFTGYLMRGYAPNRNKSWMYPPDLPYSERYPVFCSTGYVFGDLAE 319
```



```
QY 198 QYVNVSKSVPIKLEDFVFGVGLCLERLNIRLEELHLSQTPFPGLRFSVCLFRFRIVACHFI 257
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 320 KIFKVSIGIRRLHLEDVYVIGICLAKLRVDPVPPNEFVFNHRVSYSSCKYSHLITSHQF 379

QY 258 KPRTLDDYQWALENSRGDC 277
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 380 QPSELIKYNNHLOONKHAC 399

RESULT 11
QY1VE9 PRELIMINARY; PRT: 409 AA.
AC QY1VE9;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase II (Fragment).
GN B3GT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN-BLG2/MSF, AND NJL/MSF;
RA Liu Y., Kitano T., Koide T., Shirosaki T., Moriaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
   Five Mus musculus subspecies."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB039146; BAB68670.1; -
DR EMBL; AB039152; BAB68676.1; -
DR InterPro: IPR002659; Galactosyl_T.
DR Pfam: PF01762; Galactosyl_T.
DR Glycosyltransferase; Transferase.
KW NON_TER 1
FT NON_TER 409
SQ SEQUENCE 409 AA; 47196 MW; 06D262B08B0B90F4 CRC64;

Query Match 37.9%; Score 567; DB 11; Length 409;
Best Local Similarity 41.2%; Pred. No. 2.4e-46;
Matches 107; Conservative 53; Mismatches 96; Indels 4; Gaps 2;

QY 22 CROTTPFLVLLTSSHKQLAERMAIRQTGKRMVKGKLTFFLGTT--SSAAETKEV 79
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 140 COEKSPFLILLTAAEPQIEARRAIRQTGNETLAPGIQIRVFLGSIKNGYLQHA 199

QY 80 QDESQRHGDIIQKDFLDVYVNLTKTMGMIEWVHRFCPOAAFYMKTDSDMFINDVYTEL 139
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 200 QEESROYHDIQOEYLDYVNLTKTMGMNVVATYCPHTPYVVKTDSDMFVNTYLIHK 259

QY 140 LLKKNRTTR--FTGFLKLNIEPIQPFKSWFVSKSEYVNDYPPFCSGTGYVFGDVAS 197
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 260 LKPDLPPIRHNYFTGYLMRGYAPNRNKKDQWYMPDLYPSERYPVFCSGTGYVFGDLAE 319

QY 198 QYVNVSKSVPIKLEDFVFGVGLCLERLNIRLEELHLSQTPFPGLRFSVCLFRFRIVACHFI 257
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 320 KIFKVSIGIRRLHLEDVYVIGICLAKLRVDPVPPNEFVFNHRVSYSSCKYSHLITSHQF 379

QY 258 KPRTLDDYQWALENSRGDC 277
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 380 QPSELIKYNNHLOONKHAC 399

RESULT 13
QY1V19 PRELIMINARY; PRT: 409 AA.
AC QY1V19;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase II (Fragment).
GN B3GT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN-HSM/MSF, AND SWN/MSF;
RA Liu Y., Kitano T., Koide T., Shirosaki T., Moriaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
   Five Mus musculus subspecies."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB039148; BAB68672.1; -
DR EMBL; AB039150; BAB68674.1; -
DR InterPro: IPR002659; Galactosyl_T.
DR Pfam: PF01762; Galactosyl_T.
DR Glycosyltransferase; Transferase.
KW NON_TER 1
FT NON_TER 409
SQ SEQUENCE 409 AA; 47206 MW; D7BD514EE40A0157 CRC64;
```

```
QY 198 QYVNVSKSVPIKLEDFVFGVGLCLERLNIRLEELHLSQTPFPGLRFSVCLFRFRIVACHFI 257
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 320 KIFKVSIGIRRLHLEDVYVIGICLAKLRVDPVPPNEFVFNHRVSYSSCKYSHLITSHQF 379

QY 258 KPRTLDDYQWALENSRGDC 277
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 380 QPSELIKYNNHLOONKHAC 399

RESULT 11
QY1VE9 PRELIMINARY; PRT: 409 AA.
AC QY1VE9;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase II (Fragment).
GN B3GT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN-BLG2/MSF, AND NJL/MSF;
RA Liu Y., Kitano T., Koide T., Shirosaki T., Moriaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
   Five Mus musculus subspecies."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB039146; BAB68670.1; -
DR EMBL; AB039152; BAB68676.1; -
DR InterPro: IPR002659; Galactosyl_T.
DR Pfam: PF01762; Galactosyl_T.
DR Glycosyltransferase; Transferase.
KW NON_TER 1
FT NON_TER 409
SQ SEQUENCE 409 AA; 47196 MW; 06D262B08B0B90F4 CRC64;

Query Match 37.9%; Score 567; DB 11; Length 409;
Best Local Similarity 41.2%; Pred. No. 2.4e-46;
Matches 107; Conservative 53; Mismatches 96; Indels 4; Gaps 2;

QY 22 CROTTPFLVLLTSSHKQLAERMAIRQTGKRMVKGKLTFFLGTT--SSAAETKEV 79
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 140 COEKSPFLILLTAAEPQIEARRAIRQTGNETLAPGIQIRVFLGSIKNGYLQHA 199

QY 80 QDESQRHGDIIQKDFLDVYVNLTKTMGMIEWVHRFCPOAAFYMKTDSDMFINDVYTEL 139
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 200 QEESROYHDIQOEYLDYVNLTKTMGMNVVATYCPHTPYVVKTDSDMFVNTYLIHK 259

QY 140 LLKKNRTTR--FTGFLKLNIEPIQPFKSWFVSKSEYVNDYPPFCSGTGYVFGDVAS 197
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 260 LKPDLPPIRHNYFTGYLMRGYAPNRNKKDQWYMPDLYPSERYPVFCSGTGYVFGDLAE 319

QY 198 QYVNVSKSVPIKLEDFVFGVGLCLERLNIRLEELHLSQTPFPGLRFSVCLFRFRIVACHFI 257
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 320 KIFKVSIGIRRLHLEDVYVIGICLAKLRVDPVPPNEFVFNHRVSYSSCKYSHLITSHQF 379

QY 258 KPRTLDDYQWALENSRGDC 277
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 380 QPSELIKYNNHLOONKHAC 399

RESULT 12
QY1V58 PRELIMINARY; PRT: 409 AA.
AC QY1V58;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase II (Fragment).
GN B3GT2.
OS Mus musculus (Mouse).
```


| | | | | | |
|----|-----|----------------|--|---------------------------------|-----|
| Db | 266 | LLKPDLP | PRHNYFTGYLMRGYAPNRNKDSKWYMPDLYPSERYPVFCSGTGYVFSGLAE | 325 | |
| Qy | 198 | QVYVWSKSVPIKLE | DYVGLCLERLNI | LEELHSQPTFFPGGLRFSVCLFRRIVACHFI | 257 |
| Db | 326 | KIFKVS | LGIRLHLEDYVVGICLAKLRIDPVPPNEFVFNHWRVSYSSCKYSHLITSHQF | 385 | |
| Qy | 258 | KPRTL | DYWQALENS | RGEDC | 277 |
| Db | 386 | QPS | ELIKYWNHLQONKH | NAC | 405 |

Search completed: April 4, 2003, 14:45:05
Job time : 35 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2003, 14:43:40 ; Search time 15 Seconds
(without alignments)
549.228 Million cell updates/sec

Title: US-09-914-152-1_COPY_31_310
Perfect score: 1498
Sequence: 1 FKQSFVYKDGKFLKLPDT.....TLIDYQALNSRGDCPPV 280

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------------|--------------------|
| 1 | 577 | 38.5 | 326 | US-09-055-097-6 | Sequence 6, Appli |
| 2 | 387.5 | 25.9 | 378 | US-09-055-097-1 | Sequence 1, Appli |
| 3 | 357 | 23.8 | 378 | US-09-482-180A-2 | Sequence 2, Appli |
| 4 | 340 | 22.7 | 325 | US-09-055-097-5 | Sequence 5, Appli |
| 5 | 314 | 21.0 | 397 | US-09-459-133-2 | Sequence 2, Appli |
| 6 | 305 | 20.4 | 389 | US-09-459-133-13 | Sequence 13, Appli |
| 7 | 138 | 9.2 | 372 | US-08-207-904-10 | Sequence 10, Appli |
| 8 | 81.5 | 5.4 | 1513 | PCT-US93-03076-2 | Sequence 2, Appli |
| 9 | 78.5 | 5.3 | 612 | US-09-212-971-14 | Sequence 14, Appli |
| 10 | 79.5 | 5.3 | 612 | US-08-800-929A-14 | Sequence 14, Appli |
| 11 | 79.5 | 5.3 | 612 | US-08-569-749-14 | Sequence 14, Appli |
| 12 | 79.5 | 5.3 | 612 | US-09-617-053A-14 | Sequence 14, Appli |
| 13 | 79.5 | 5.3 | 612 | PCT-US96-12860-14 | Sequence 14, Appli |
| 14 | 78 | 5.2 | 438 | PCT-US95-05922A-2 | Sequence 2, Appli |
| 15 | 78 | 5.2 | 618 | US-08-569-749-2 | Sequence 2, Appli |
| 16 | 78 | 5.2 | 618 | US-09-069-023-29 | Sequence 29, Appli |
| 17 | 78 | 5.2 | 618 | PCT-US96-12860-2 | Sequence 2, Appli |
| 18 | 78 | 5.2 | 740 | US-08-276-968A-18 | Sequence 18, Appli |
| 19 | 77 | 5.1 | 741 | US-08-276-968A-20 | Sequence 20, Appli |
| 20 | 76 | 5.1 | 454 | US-08-276-968A-22 | Sequence 22, Appli |
| 21 | 76 | 5.1 | 923 | US-09-397-885-1 | Sequence 1, Appli |
| 22 | 75.5 | 5.0 | 331 | US-09-134-001C-3626 | Sequence 3626, Ap |
| 23 | 75.5 | 5.0 | 821 | US-09-377-465A-2 | Sequence 2, Appli |
| 24 | 75 | 5.0 | 740 | US-08-016-863-16 | Sequence 16, Appli |
| 25 | 75 | 5.0 | 740 | US-08-276-968A-16 | Sequence 16, Appli |
| 26 | 73.5 | 4.9 | 329 | US-09-651-200-19 | Sequence 19, Appli |
| 27 | 73.5 | 4.9 | 340 | US-09-651-200-2 | Sequence 2, Appli |

Query Match 38.5% Score 577: DB 2: Length 326:

| | | | | | | |
|----|------|-----|------|---|--------------------|--------------------|
| 28 | 73.5 | 4.9 | 441 | 4 | US-09-651-200-4 | Sequence 4, Appli |
| 29 | 73.5 | 4.9 | 534 | 4 | US-09-651-200-6 | Sequence 6, Appli |
| 30 | 73.5 | 4.9 | 534 | 4 | US-09-651-200-24 | Sequence 24, Appli |
| 31 | 73.5 | 4.9 | 802 | 4 | US-09-433-043B-120 | Sequence 120, App |
| 32 | 73.5 | 4.9 | 2237 | 1 | US-08-354-973-1 | Sequence 1, Appli |
| 33 | 73.5 | 4.9 | 3177 | 2 | US-08-477-451-4 | Sequence 4, Appli |
| 34 | 72.5 | 4.8 | 289 | 1 | US-07-593-657-7 | Sequence 7, Appli |
| 35 | 72.5 | 4.8 | 289 | 4 | US-08-942-012B-4 | Sequence 4, Appli |
| 36 | 72.5 | 4.8 | 517 | 1 | US-08-457-274A-2 | Sequence 2, Appli |
| 37 | 72.5 | 4.8 | 517 | 5 | PCT-US95-05758-2 | Sequence 2, Appli |
| 38 | 72 | 4.8 | 363 | 3 | US-08-586-165-7 | Sequence 7, Appli |
| 39 | 72 | 4.8 | 732 | 1 | US-08-481-626-2 | Sequence 2, Appli |
| 40 | 72 | 4.8 | 732 | 4 | US-08-989-299-4 | Sequence 4, Appli |
| 41 | 72 | 4.8 | 887 | 1 | US-08-327-494A-4 | Sequence 4, Appli |
| 42 | 72 | 4.8 | 887 | 5 | PCT-US95-13659-4 | Sequence 4, Appli |
| 43 | 72 | 4.8 | 1306 | 4 | US-08-989-299-7 | Sequence 7, Appli |
| 44 | 71.5 | 4.8 | 417 | 1 | US-08-351-981-7 | Sequence 7, Appli |
| 45 | 71.5 | 4.8 | 517 | 1 | US-08-457-274A-28 | Sequence 28, Appli |

ALIGNMENTS

RESULT 1
US-09-055-097-6
; Sequence 6, Application US/09055097
; Patent No. 5955282
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; APPLICANT: Patterson, Chandra
; TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/055,097
; FILING DATE: Filed Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Carrone, Michael C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0490 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 326 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 2745735
US-09-055-097-6

```
Best Local Similarity 42.8%; Score 1.9e-61;
Matches 116; Conservative 46; Mismatches 93; Indels 16; Gaps 4;

Qy 14 FLKLPDTCRTPPLVLLVSSHQKLAERMAIROTWGKERWVGKQKLTFFLLGTTSSA 73
Db 65 FLINPNKCKENIPFLVILSTHKEFDARQAIRETGDENFKGKIATLFLGKNADP 124
Qy 74 AETVEDQESORHGDIQKDFLDVYNNLTLMGMGIEWHRCFPOAAFMKTDSDFINV 133
Db 125 VLNQWVEQESIFDIIVEIDSDVSHNLTLMGMRVAVTECSKAKYMKTDSDFINM 184
Qy 134 DYLTELLKKNR--TTREFTGFLKNEPPIROPFSKVFVSKSEYPDWRYPPFCSTGVXF 191
Db 185 DNLVYKLLKPKSTKRRRYFTGYV-INGGPIRQVRSKWTMPRDLYPDSNYPFCSTGVIF 243
Qy 192 SGDVASQVYNNVSKSVPIKLEDFVGLCLERLNIRLEELHSHQTPFFPGG----LRFVVC 246
Db 244 SADVAELIKYKLSLTRLUHLSDVYVGLCKRLGHI-----PFQNGFNHMKWAYSLC 295
Qy 247 LFRIVACHFIKPRITLLDYWQALENSRGEDC 277
Db 296 RYRRVITVHOISPEMHRIWDMSSKKHLRC 326

RESULT 2
US-09-055-097-1
; Sequence 1, Application US/09055097
; Patent No. 5955282
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; APPLICANT: Patterson, Chandra
; TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA: US/09/055,097
; FILING DATE: Filed Herewith
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0490 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 853-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 378 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: DUODN0702
; CLONE: 1705085
; US-09-055-097-1
```

```
Query Match 25.9%; Score 387.5; DB 2; Length 378;
Best Local Similarity 30.2%; Pred. No. 2.4e-38;
Matches 88; Conservative 57; Mismatches 105; Indels 41; Gaps 6;

Qy 26 PPFLVLLVSSHQKLAERMAIROTWGKERWVGKQKLTFFLLGTTSS-----AAETKEYD 80
Db 70 PPFLVLLVSSHQKLAERMAIROTWGKERWVGKQKLTFFLLGTTSS-----AAETKEYD 80
Qy 81 QESORHGDIQKDFLDVYNNLTLMGMGIEWHRCFPOAAFMKTDSDFINV-DYLTTEL 139
Db 130 SESAAOGDILQAAFDOSYRNLTLMGSLNWAECHEPMARYVLTDDDDVYVNPVELVSEL 189
Qy 140 LKKNRTTRF-----FTGFLKNEPPIROPFSKVFVSKS 173
Db 190 VLRGGRWQWERSTPEQAEQEGOVHLSEVPLLYLGRVHWRVNPRTPGRRHRVSE 249
Qy 174 EYP--WDYRPPPCSTGYVFGDVASQVYNNVSKSVPIKLEDFVGLCLERLNIRLEELH 231
Db 250 QMPHTWGPFPYVSGTYVLSASAVQLILKVASRAPLLPLEDFVGVASARRGL----AP 305
Qy 232 SQPTFFPGGLRFV---CLFRRIVACHFIKPRITLLDYWQALENSRGEDC 279
Db 306 TQCVKLAGATHYDRCCYGRKELTSHRLDPWKMQEAAMKLVGSGDERTAP 356

RESULT 3
US-09-482-180A-2
; Sequence 2, Application US/09482180A
; Patent No. 6361985
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Yamamoto, Gayle
; APPLICANT: Gao, Zeren
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Jaspers, Stephen
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING
; FILE REFERENCE: BETA-1.3-GALACTOSYLTRANSFERASE HOMOLOG, ZN5SP6
; CURRENT APPLICATION NUMBER: US/09/482,180A
; CURRENT FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/115,721
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-482-180A-2

Query Match 23.8%; Score 357; DB 4; Length 378;
Best Local Similarity 32.6%; Pred. No. 1.2e-34;
Matches 90; Conservative 42; Mismatches 134; Indels 10; Gaps 5;

Qy 8 YKKGDNFLKLPDTCRTPPLVLLVSSHQKLAERMAIROTWGK-ERMVKGKQLKTF 66
Db 99 YRHCRNFSILPEPSGCSKDTFLLLAIKSGQGHVERRAIRSTWGRVGGWARGRLKLVFL 158
Qy 67 LGTTSSAAETKEVDQESORHGDIQKDFLDVYNNLTLMGMGIEWHRCFPOAAFMKT 126
Db 159 LGVAGSAPPAOLLAYESREFDDILQMDTEDFNLTLELHLQRVVVAACPOAHFMLKGD 218
Qy 127 SDMFINDVLTLELLKKNRTTRFTFGFLKNEPPIROPFSKVFVSKSEYPMWRYPFC 186
Db 219 DDVFWHVPNVLEFDGWDPAODLLVGDVIRQALPNRNTKVKYFIPPSMYRATHYPYAGG 278
Qy 187 TGYVFGDVASQVYNNVSKSVPIKLEDFVGLCLERLNIRLEELHSHQTPFFPGGLR---- 242
Db 279 GGYVMSRATVRRQLQAIMEDAELLSIDDFVGVGMLRLG--LSPMH-HAGFKTFGIRRLD 335
Qy 243 -FSVCLFRIVACHFIKPRITLLDYWQALENSRGEDC 277
Db 336 PLDPCLYRGLLVHRLSPLEMTMW-ALVTDGLKLC 370
```

```

Db 297 YSSVIASHEFGDPEENTRW 316

RESULT 5
US-09-459-133-2
; Sequence 2, Application US/09459133
; Patent No. 6416988
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Yamamoto, Gayle
; APPLICANT: Jaspers, Stephen R.
; APPLICANT: Gao, Zeren
; TITLE OF INVENTION: BETA-1,3-GALACTOSYLTRANSFERASE HOMOLOGS
; FILE REFERENCE: 98-77
; CURRENT APPLICATION NUMBER: US/09/459,133
; CURRENT FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/111,697
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (137)...(137)
; OTHER INFORMATION: Xaa is Gly or Ser
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(397)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-459-133-2

Query Match 21.08; Score 314; DB 4; Length 39;
Best Local Similarity 29.98; Pred. No. 2.2e-29;
Matches 78; Conservative 50; Mismatches 117; Indels

Qy 19 DTCROTQPPFLVLVLTSSHQLAERMAITROTWGERVMVKQKLTFFLLGTC
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 144 DTDV-----PYLLAVKSEPGFAERQAVRTGWS----PAGIRLLFLGSP
Qy 79 VDQ-----ESQRGDIIOKDFLDVYNI.LTKTMGIEWHRCFQQAFAVMKTD
:| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 195 LDSLVAVESRRYSDLLWDFLDVPFNQTLKDLLLAWLGRHCPVTSFVLRAQ
Qy 135 YLTELL--LKKNRTTFFTCFLKNEEPIRQPSKWFVSKSEYPWDRYPPFC
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 255 ALLAHLRALPASARSYLGEVFTQAMPLRKPGGPFYVPESFFE-GGYPA
Qy 193 GDVASQVYNNVSKSPYIKLDEYFVGLCLERLNI.RLELHSQPTFFPGGLRFS
| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 314 GRLPWLLRAARVAPFFEDVYTGICIRGLVPOAHPGLTANPADRTAD
Qy 253 ACHFIKPTLLDYWQALENSR 273
:| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 374 LVRLPGPQASIRLWKQLQDPR 394

RESULT 6
US-09-459-133-13
; Sequence 13, Application US/09459133
; Patent No. 6416988
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Yamamoto, Gayle
; APPLICANT: Jaspers, Stephen R.
; APPLICANT: Gao, Zeren
; TITLE OF INVENTION: BETA-1,3-GALACTOSYLTRANSFERASE HOMOLOGS
; FILE REFERENCE: 98-77
; CURRENT APPLICATION NUMBER: US/09/459,133
; CURRENT FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/111,697

```

```
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-459-133-13

Query Match          20.4%; Score 305; DB 4; Length 389;
Best Local Similarity 31.0%; Pred. No. 2.6e-28;
Matches 77; Conservative 41; Mismatches 120; Indels 10; Gaps 5;

Qy 27 PFLVLLTSSHKQLAERMAIROTGWKERMVKGKQLTKTFFLLGTTSSAAE---TKVEDQES 83
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 140 PYLLAVKSEPHFAARQAVRETWGP--VAGTRL--LFLLLGSLMGSGPDLRLSLVTWES 195

Qy 84 QRHGDIIOKDFLDVYNNLTLTMTMGIEVHRCFQAAAFVMTKDSDFINVDYLTLL--L 141
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 196 RRYGDLMLWDLDVFNRTLDKLLLTWLSHCHCPDVFVQLQDDAFHHPALLEHLOTL 255

Qy 142 KKNRTTFFGTGFLKLNFPPIQPFKSWFVSKSEYPMWRYPFPCSGTGVVFGSDVASQVYN 201
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 256 PPTWARSYLGEIFTQAKPLKPGPFYVPTFFEGD-YPAYASGGVVISGRAPWLLQ 314

Qy 202 VSKSVYPYIKLEDFVGLCLERINIRLEHLSOPTFFPGGLRFSVCLFRIRIVACHFIKPR 261
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 315 AAARVAPFPDDVYTGFCFRALGLAPRAHPGFLTAWPAERTRDPCAVRGLLLVHPVSPD 374

Qy 262 LLDYHQAL 269
   : : : : :
Db 375 TIWLRRHL 382

RESULT 7
US-08-207-904-10
; Sequence 10, Application US/08207904
; Patent No. 5477002
; GENERAL INFORMATION:
; APPLICANT: Tuttle, AnnMarie
; APPLICANT: Crossland, Lyle D.
; TITLE OF INVENTION: Anther-Specific cDNA Sequences, Genomic
; TITLE OF INVENTION: DNA Sequences and Recombinant DNA Sequences
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/207.904
; CLASSIFICATION: 800
; PRIOR APPLICATION NUMBER: US/07/908.242
; APPLICATION NUMBER:
; FILING DATE:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: CCC 1624
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8615
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
```

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-207-904-10

Query Match          9.2%; Score 138; DB 1; Length 372;
Best Local Similarity 23.7%; Pred. No. 4.7e-08;
Matches 53; Conservative 41; Mismatches 86; Indels 44; Gaps 9;

Qy 28 FLVLLVTSSSHKQLAERMAIROTGWKERMVKGKQLKTF-----FLLG--TTSSAAET 76
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 103 FIVVGINTAFSSRRKRDLSRETW---MPKDGKLRKLEKEKIVIRFVIGHSATRGVLD 158

Qy 77 KEVDQESQRHGDIIOKDFLDVYNNLTLTMTMGIE-----WVHRCFQAAAFVMTKDSDFI 131
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 159 RAIDSEDAQYKDFLRLDHVEGYHELSTKRLYFSKAVSIW-----DADFVYKVDVDDL 212

Qy 132 NVDYITELLKKNRTTRFTGTGLK----LNEFPPIQPFKSWFVSKSEYPM-----DRYP 181
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 213 NLGMLANTLAKYKSRPVYIGCMKSGPVLSQKGR-----YYEPEY-WKEFEGGNKYF 264

Qy 182 PFCSGTGVVFGSDVASQVYNVSKSVPIKLEDFVGLCLERLNI 225
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 265 RHATQIYIGISRDLSYISINSILHRYANEDVSLGSLIGLEV 308

RESULT 8
PCT-US93-03076-2
; Sequence 2, Application PC/TUS9303076
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; TITLE OF INVENTION: GAP-Associated Protein p190 and
; TITLE OF INVENTION: Transduction
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: 2 Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/03076
; FILING DATE: 19930331
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WHI92-03A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1513 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US93-03076-2

Query Match          5.4%; Score 81.5; DB 5; Length 1513;
Best Local Similarity 23.6%; Pred. No. 3;
Matches 37; Conservative 22; Mismatches 63; Indels 35; Gaps 5;

Qy 42 ERMAIROTGWKERMVKGKQLTKTFFLLGTTSSAAETKEVDQESQRHGDIIOKDFLDVYNNL 101
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 137 DOLGLEQDFEQKMPDGLKLLVDGFLLGIDVSRGMNRNFD-----QLAFVSNLYNQ 187
```

```

: APPLICANT: Baird, Stephen
: APPLICANT: Tsang, Benjamin K
: APPLICANT: Pratt, Christine
: TITLE OF INVENTION: DETECTION AND MODULATION OF
: TITLE OF INVENTION: IAPS AND NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFER
: TITLE OF INVENTION: DISEASE
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Clark & Elbing LLP
: STREET: 176 Federal Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ FOR Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/800,929A
: FILING DATE: 13-FEB-1997
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/030,590
: FILING DATE: 14-NOV-1996
: APPLICATION NUMBER: 60/017,354
: FILING DATE: 26-APR-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Bleker-Brady, Kristina
: REGISTRATION NUMBER:
: REFERENCE/DOCKET NUMBER: 07891/009001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-428-0200
: TELEFAX: 617-428-7045
: TELEX:
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 612 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-800-929A-14

Query Match          5.3%; Score 79.5; DB 4; Length 612;
Best Local Similarity 19.9%; Pred. No. 1.3;
Matches 50; Conservative 35; Mismatches 79; Indels 87; Gaps 10;

QY 29 LVLVLTSSHKQLAE-----RMAIROTWGKERVMVKGKQLKT-----FFILGTTSSAAETKE 78
Db 382 VMMSTPVVKAALMGFSRSRVQRTVQRQILATGENTYRTVNDIVSVLLNAEDEREEKE 441
QY 79 VDQESQRHGD--IIOKFDLVVYNYLT-----LKTMMGIEWVHRFCPCQAAYVMKTDSDMF 131
Db 442 RQTEMASGDSLRLKRNMAFLPQOLTHVLPILDNL-----EASVITKQEHDIIR 491
QY 132 NV-----DYLTELLKKNRTTFFTGFLKLNKEFFIRQPFSEKWFVSKSEYPMWDRYPPF 183
Db 492 KQTOIPLQARELIDTVLVKGNAANIFKNSLK-----SGLSLEEQRLRRLOEERT----- 523
QY 184 CSGTGYVFGSDVASQVYN---VKSVPYIKLEDVFVGLCLERLNIRBELHSQTPFPFGG 240
Db 524 -----EIDSTLYENLFVEKNMKYIPTEDV-SGLSLEEQRLRRLOEERT----- 564
QY 241 LRFSVGLFRRI 251
Db 565 --CKVCMDEV 573

RESULT 11
US-08-569-749-14
Sequence 14 Application US/08569740

```


Patent No. 6187557
; GENERAL INFORMATION:
; APPLICANT: Rothe, Mike
; APPLICANT: Goeddel, David V
; TITLE OF INVENTION: INHIBITORS OF APOPTOSIS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/569,749
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Brezner, David J.
; REGISTRATION NUMBER: 24,774
; REFERENCE/DOCKET NUMBER: A-62464/DJB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)781-1989
; TELEFAX: (415)398-3249
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 612 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-569-749-14

Query Match 5.3%; Score 79.5; DB 4; Length 612;
Best Local Similarity 19.9%; Pred. No. 1.3;
Matches 50; Conservative 35; Mismatches 79; Indels 87; Gaps 10;

QY 29 LVLLVTSSSHQKLAEE---RMAIRQWGWKRMVKGKOLKT-----FFLLGTTSSAAETKE 78
Db 382 VYMMSTPVVKAALMGFSSRLVQRTVQRIATGENTYVNDIVSVLLNAEDERREEKE 441
QY 79 VQESORHGD--IIQKDFLDVYVNLT-----LKTMMGIEWVHRCFQAAAFVVKMTSDMFI 131
Db 442 ROTEEMASGDLIRKRNMALEFQQLTHVLPILDNL-----EASVITKQEHDIIR 491
QY 132 NV-----DYLTLLKKNRTTFFTGFLKLNFEPIRQPFKSWFVSKSEYPWDYPPF 183
Db 492 QKTQIPLOARELIDTVLVKGNAANIFKNSLK-----523
QY 184 CSGTGVFSGDVASOVYN---VSKSVPIKLEDFVGLCLERLNIRLELHLSQPTFFPGG 240
Db 524 -----EIDSTLYENLFVEKNMKYIPTEDV-SGLSLEQLRLQERT-----564
QY 241 LRFVSVCLFRRI 251
Db 565 --CKVCMDEV 573

RESULT 12
US-09-617-053A-14
; Sequence 14, Application US/09617053A
; Patent No. 6300492
; GENERAL INFORMATION:
; APPLICANT: Korneiluk, Robert G
; APPLICANT: Mackenzie, Alexander E
; APPLICANT: Liston, Peter
; APPLICANT: Baird, Stephen
; APPLICANT: Tsang, Benjamin K

; APPLICANT: Pratt, Christine
; TITLE OF INVENTION: DETECTION AND MODULATION OF IAPS AND
; TITLE OF INVENTION: NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE
; TITLE OF INVENTION: DISEASE
; FILE REFERENCE: 07891/009003
; CURRENT APPLICATION NUMBER: US/09/617,053A
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/800,929
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 612
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-617-053A-14

Query Match 5.3%; Score 79.5; DB 4; Length 612;
Best Local Similarity 19.9%; Pred. No. 1.3;
Matches 50; Conservative 35; Mismatches 79; Indels 87; Gaps 10;

QY 29 LVLLVTSSSHQKLAEE---RMAIRQWGWKRMVKGKOLKT-----FFLLGTTSSAAETKE 78
Db 382 VYMMSTPVVKAALMGFSSRLVQRTVQRIATGENTYVNDIVSVLLNAEDERREEKE 441
QY 79 VQESORHGD--IIQKDFLDVYVNLT-----LKTMMGIEWVHRCFQAAAFVVKMTSDMFI 131
Db 442 ROTEEMASGDLIRKRNMALEFQQLTHVLPILDNL-----EASVITKQEHDIIR 491
QY 132 NV-----DYLTLLKKNRTTFFTGFLKLNFEPIRQPFKSWFVSKSEYPWDYPPF 183
Db 492 QKTQIPLOARELIDTVLVKGNAANIFKNSLK-----523
QY 184 CSGTGVFSGDVASOVYN---VSKSVPIKLEDFVGLCLERLNIRLELHLSQPTFFPGG 240
Db 524 -----EIDSTLYENLFVEKNMKYIPTEDV-SGLSLEQLRLQERT-----564
QY 241 LRFVSVCLFRRI 251
Db 565 --CKVCMDEV 573

RESULT 13
PCT-US96-12860-14
; Sequence 14, Application PC/TUS9612860
; GENERAL INFORMATION:
; APPLICANT: TULARIK, INC.
; TITLE OF INVENTION: INHIBITORS OF APOPTOSIS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/12860
; FILING DATE: 06 AUG 1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Serial Nos. 08/512,946 & 08/569,749
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brezner, David J.
; REGISTRATION NUMBER: 24,774
; REFERENCE/DOCKET NUMBER: A-62464/DJB
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (415)781-1989
TELEFAX: (415)398-3249
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 612 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-12860-14

Query Match 5.3%; Score 79.5; DB 5; Length 612;
Best Local Similarity 19.9%; Pred. No. 1.3;
Matches 50; Conservative 35; Mismatches 79; Indels 87; Gaps 10;

QY 29 LVLVTSKSHKOLAE---RMAIROTGWKERMVKGKOLKT-----PFLGLTSSAAETKE 78
DB 382 VVMSTPVVAAALMGFSRLVROTQVQIILATGENTRTVNDIVSVLNAEDERREEKE 441
QY 79 VDQESORHGD--IIOKDFLDVYVNYLT-----LKTMMGIEWHRCFQAAAFVMTKDSDFI 131
DB 442 RQTEEMASGDLSLIRKNRMALFQOLTHVLPILDNL-----EASVITKQEHDIIR 491
QY 132 NV-----DYTELKLNKTRTFTGFLKLNFFPIQPFKSWFVSKSEYPMDRYPPF 183
DB 492 QKTQIPLQARELIDTVLVKGNAAANIFKNSLK----- 523
QY 184 CSGTGYVFGDVASQVYN---VSKSVPIYKLEDFVGLCLERLNIRLEEHSQTFPPGG 240
DB 524 -----EIDSTLYENLFVKNMKIPTEDV-SGLSLEEQRLRLQERT----- 564
QY 241 LRFVSVCLFRRI 251
DB 565 --CKVCMDEV 573

RESULT 14
PCT-US95-05922A-2
Sequence 2, Application PC/TUS9505922A
GENERAL INFORMATION:
APPLICANT: HE, ET AL.
TITLE OF INVENTION: Human Inhibitor of Apoptosis Gene 1
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05922A
FILING DATE: 11 MAY 1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-292
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 438 AMINO ACIDS

TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
PCT-US95-05922A-2
Query Match 5.2%; Score 78; DB 5; Length 438;
Best Local Similarity 19.3%; Pred. No. 1.2;
Matches 52; Conservative 46; Mismatches 88; Indels 84; Gaps 11;
QY 2 KEQSFVYKKGDNF-----LKLPTDTCRQT--PPFL-----VLLVTSKSHKOL 40
DB 160 KGQEFVDEIQRYPHLLLEQLLSTSDTTGEENADPPILHFGPGCESSEDAVMMNTPWAKSA 219
QY 41 AE---RMAIROTGWKERMVKGKOLKTFFLLCTTSSAAETKEVDQESORHGD----- 88
DB 220 LEMGFNDLVKQTVQSKILTTGENTRTVNDIVSALLNAEDEKREEKEQAEEMASDDLS 279
QY 89 IIOKDFLDVYVNYLT-----LKTMMGIEWHRCFQAAAFVMTKDSDFI INV-DYLTLLK 142
DB 280 LIRKNRMALFOQLTCVLPILDNLKANVINK---QEHDIIKQKTOIPLQARELIDTILVK 336
QY 143 KNRTTRFTGFLKLNFFPIQPFKSWFVSKSEYPMDRYPPFCSTGYVFGDVASQVYN- 201
DB 337 GNAANIFKNSLK-----EIDSTLYKN 358
QY 202 --VSKSVPIYKLEDFVGLCLERLNIRLEE 229
DB 359 LFVDKMKKIPTEDV-SGLSLEEQRLRLQ 387

RESULT 15
US-08-569-749-2
Sequence 2, Application US/08569749
Patent No. 6187557
GENERAL INFORMATION:
APPLICANT: Rothe, Mike
TITLE OF INVENTION: INHIBITORS OF APOPTOSIS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/569,749
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Brezner, David J.
REGISTRATION NUMBER: 24,774
REFERENCE/DOCKET NUMBER: A-62464/DJB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)781-1989
TELEFAX: (415)398-3249
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 618 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-569-749-2
Query Match 5.2%; Score 78; DB 4; Length 618;

Best Local Similarity 19.3%; Pred. No. 2;
Matches 52; Conservative 46; Mismatches 88; Indels 84; Gaps 11;

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

2

340

41

400

89

460

143

517

202

539

KEQSFVYKKDGNF-----LKLPTDCRQT--PPFL-----VLLVTSSHKQL 40

KGOEFVDEIQGRYPHLLLEQLLSTSDTTGGENADPPIIHFGPGESSEDAVMMNTPVVKSA 399

AE-----RMAIROTWGKRMVKQKLTFFLLGTTSSAAETKEVDQESQRHGD----- 88

LENGFNRLVKQTVQSKILTTGENYKTVNDIVSALLNAEDEKREEKEKQAEEMASDDLS 459

IIOKDFLOWVYNYLT-----LKTMMGIEWVHRECPQAAFYMKTDSDMFINV-DYLTTELLK 142

LIRKNRMALFQQLTCVLPILDNLLKANVINK---QEHDIKOKTQIPLOARELIDTILVK 516

KNRTTRFFTGFLKLNFFPIRQPFKWFVSKSEYPWDPRYPPFCSGTGYVFGDVASOVYN- 201

GNAANIFKNCUK-----EIDSTLYKN 538

--VSKSVPYIKLEDVFGLCLERLNIRLEE 229

LVVDKNMKYIPTEDV-SGLSLEEQLRRLOE 567

Search completed: April 4, 2003, 14:45:52

Job time : 18 secs

QY 192 SGDVASQVNVKSVPIKLEDFVGLCLERLNIRLEELHSOFTFFPGG-----LRFSSVC 246
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 244 SADVAELIYKTSLHTRLLHLEDVYVGLCLRLKLGIIH-----PFQNSGFNHNKWMAYSIC 295
QY 247 LERRIVACHFIKPRTLDDYVQALENSRGEDC 277
: | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 296 RYRRVITVHQISPEEMHRIWINDMSSKKHLRC 326

RESULT 2

US-09-804-006-8
; Sequence 8, Application US/09804006
; Patent No. US20020119517A1
; GENERAL INFORMATION:
; APPLICANT: White, David
; APPLICANT: Zhou, Jianghong
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: LEPTIN INDUCED GENES
; FILE REFERENCE: 07334/126001
; CURRENT APPLICATION NUMBER: US/09/804,006
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US/09/292,228
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: US 60/108,379
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: US 09/150,857
; PRIOR FILING DATE: 1998-09-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-804-006-8

Query Match 38.5%; Score 577; DB 10; Length 326;
Best Local Similarity 42.8%; Pred. No. 7.1e-50;
Matches 116; Conservative 46; Mismatches 93; Indels 16; Gaps 4;

QY 14 FLKLPTDCRQTPPFLVLLVTSKQLAERMAIRQTWGERMVKGKQLKTFLLGTTSSA 73
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 65 FLINPEKNEKPIFLVILISTHKEFDARQAIRETWGDENNFKGIKTIATFLLGKNA DP 124
QY 74 AETKEVDQESQRHGDIIQKDFLDVYVNLTKTMGIEWVHRCPCQAFAVYKMTSDMFINV 133
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 125 VLNMVQESQIFHDIIEDFDISYHNLTCLKMGRWATFCSKAKYVYKMTSDMFINV 184
QY 134 DVLTELLKKNR--TTRFTGFLKNEFPIROPFSKWFVSKSEYPMWDRYPPFCSGTGYVF 191
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 185 DNLVYLLKLPSTKPRRYFTGYV--INGPDRVRSKWYAPRDLYPDSNTPPCSGTGYIF 243
QY 192 SGDVASQVNVKSVPIKLEDFVGLCLERLNIRLEELHSOFTFFPGG-----LRFSSVC 246
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 244 SADVAELIYKTSLHTRLLHLEDVYVGLCLRLKLGIIH-----PFQNSGFNHNKWMAYSIC 295
QY 247 LERRIVACHFIKPRTLDDYVQALENSRGEDC 277
: | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 296 RYRRVITVHQISPEEMHRIWINDMSSKKHLRC 326

RESULT 3

US-09-804-357-11
; Sequence 11, Application US/09804357
; Patent No. US20010024808A1
; GENERAL INFORMATION:
; APPLICANT: White, David
; APPLICANT: Zhou, Jianghong
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: LEPTIN INDUCED GENES
; FILE REFERENCE: 07334/109001
; CURRENT APPLICATION NUMBER: US/09/804,357
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/195,896

; PRIOR FILING DATE: 1998-11-19
; PRIOR APPLICATION NUMBER: US 60/108,379
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: US 09/150,857
; PRIOR FILING DATE: 1998-09-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-804-357-11

Query Match 38.0%; Score 569; DB 10; Length 422;
Best Local Similarity 41.5%; Pred. No. 6.3e-49;
Matches 108; Conservative 53; Mismatches 95; Indels 4; Gaps 2;
QY 22 CROTTPFLVLLVTSKQLAERMAIRQTWGERMVKGKQLKTFLLGTT--SSAAETKEV 79
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 146 COEKSPFLILLIAAEPQIEARRAIRQTWGNESLAPGIIQITRIFLLGLSLKNGYLQRAI 205
QY 80 DOESQRHGDIIQKDFLDVYVNLTKTMGIEWVHRCPCQAFAVYKMTSDMFINVLYLTEL 139
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 206 LEESQYHDIIOEQEYLDIYVNLTKLMGNWVATYCPHIPYVYKMTSDMFVNTYELINK 265
QY 140 LKKNRTTR--FFTGFLLKNEFPIROPFSKWFVSKSEYPMWDRYPPFCSGTGYVFGDVAS 197
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 266 LKAPDLPPRHNYFTGYLMRGYAPNRNKKSKWYPPDLYPSERYPVFCSGTGYVFGDLAE 325
QY 198 QVYVSKSVPIKLEDFVGLCLERLNIRLEELHSOFTFFPGGLRFSVCLFRIRIVACHFI 257
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 326 KIFKVSGLIRRLHLEDVYVGLCLAKLRIDVPPNPEFVHNRVSYSSCKYSHLITSHQF 385
QY 258 KPRTLDDYQALENSRGEDC 277
: | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 386 QPSELIKYWNHLOONKHAC 405

RESULT 4

US-09-529-063-25
; Sequence 25, Application US/09529063
; Patent No. US20020102542A1
; GENERAL INFORMATION:
; APPLICANT: FUKUSHIMA, DAIKICHI
; APPLICANT: SHIBAYAMA, SHIRO
; APPLICANT: TADA, HIDEAKI
; TITLE OF INVENTION: POLYPEPTIDE, CDNA ENCODING THE POLYPEPTIDE, AND USE OF
; TITLE OF INVENTION: THE BOTH
; FILE REFERENCE: Q58769
; CURRENT APPLICATION NUMBER: US/09/529,063
; CURRENT FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: PCT/JP98/04514
; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: JP 9-274674
; PRIOR FILING DATE: 1997-10-07
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-063-25

Query Match 38.0%; Score 569; DB 10; Length 422;
Best Local Similarity 41.5%; Pred. No. 6.3e-49;
Matches 108; Conservative 53; Mismatches 95; Indels 4; Gaps 2;
QY 22 CROTTPFLVLLVTSKQLAERMAIRQTWGERMVKGKQLKTFLLGTT--SSAAETKEV 79
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 146 COEKSPFLILLIAAEPQIEARRAIRQTWGNESLAPGIIQITRIFLLGLSLKNGYLQRAI 205
QY 80 DOESQRHGDIIQKDFLDVYVNLTKTMGIEWVHRCPCQAFAVYKMTSDMFINVLYLTEL 139
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 52 FYEYPIYRQDFRFTLREHSNCSHQNPFLVLTSPSDVKARQAIQVTVTGWGKSWWGYE 111
QY 61 LKTFLLGTTSSAAE---TKEVDOSORHGDIIQKDFLDVYVNLTKTMGIEWVHRFCP 117
Db 112 VLTFFLLGQOAREDTXLSLSEDEHVLGDIIRQDFLDYNNLTUKTMAFRWMEFCP 171
QY 118 QAAFMKTDSDMFINVYDITELLKKNRTRFTFTGFLKNEFFIROPFKSWFVSKSEYWP 177
Db 172 NAKYIMKTDVDFINTGNLVKYLNLNHSKFFTGYPGLIDNYSYRGFFHKNHISYQEYFP 231
QY 178 DRYPPFCSTGYVFSGDVASQVYVSKVPYIKLEDFVGLCLERLNIRLEELHSQPTFF 237
Db 232 KVFPPYCSGLGYIMSGDLVPRVYEMSHVKPIKFEDVYVGCILNLLKVDIHIPEDTNLFF 291
QY 238 PGGLRFSVCLFRRIIVACHFIKPRTLDDYWOAL 269
Db 292 LYRIHLDVCLRRVIAAHGFSKKEIITFWQVM 323
RESULT 8
US-09-804-006-9
; Sequence 9, Application US/09804006
; Patent No. US20020119517A1
; GENERAL INFORMATION:
; APPLICANT: White, David
; APPLICANT: Zhou, Jianghong
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: LEFTIN INDUCED GENES
; FILE REFERENCE: 07334/126001
; CURRENT APPLICATION NUMBER: US/09/804,006
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/292,228
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: US 60/108,379
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: US 09/150,857
; PRIOR FILING DATE: 1998-09-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-804-006-9
Query Match 37.1%; Score 555.5; DB 10; Length 331;
Best Local Similarity 39.3%; Pred. No. 1e-47;
Matches 107; Conservative 51; Mismatches 111; Indels 3; Gaps 1;
QY 1 FKEQSFVYKKGDFLKLPTDCRQTPPFLVLTSSHKQLAERMAIROTGWKERMVKGKQ 60
Db 52 FYEYPIYRQDFRFTLREHSNCSHQNPFLVLTSPSDVKARQAIQVTVTGWGKSWWGYE 111
QY 61 LKTFLLGTTSSAAE---TKEVDOSORHGDIIQKDFLDVYVNLTKTMGIEWVHRFCP 117
Db 112 VLTFFLLGQOAREDTXLSLSEDEHVLGDIIRQDFLDYNNLTUKTMAFRWMEFCP 171
QY 118 QAAFMKTDSDMFINVYDITELLKKNRTRFTFTGFLKNEFFIROPFKSWFVSKSEYWP 177
Db 172 NAKYIMKTDVDFINTGNLVKYLNLNHSKFFTGYPGLIDNYSYRGFFHKNHISYQEYFP 231
QY 178 DRYPPFCSTGYVFSGDVASQVYVSKVPYIKLEDFVGLCLERLNIRLEELHSQPTFF 237
Db 232 KVFPPYCSGLGYIMSGDLVPRVYEMSHVKPIKFEDVYVGCILNLLKVDIHIPEDTNLFF 291
QY 238 PGGLRFSVCLFRRIIVACHFIKPRTLDDYWOAL 269
Db 292 LYRIHLDVCLRRVIAAHGFSKKEIITFWQVM 323

RESULT 9
US-09-739-451-2
; Sequence 2, Application US/09739451

; Patent No. US20010024813A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Dendriac and Brainiac-3
; FILE REFERENCE: PF464
; CURRENT APPLICATION NUMBER: US/09/739,451
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/213,364
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: 60/077,687
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/108,928
; PRIOR FILING DATE: 1998-11-17
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-739-451-2
Query Match 36.6%; Score 548.5; DB 10; Length 319;
Best Local Similarity 38.6%; Pred. No. 4.9e-47;
Matches 105; Conservative 53; Mismatches 111; Indels 3; Gaps 1;
QY 1 FKEQSFVYKKGDFLKLPTDCRQTPPFLVLTSSHKQLAERMAIROTGWKERMVKGKQ 60
Db 40 FYEYPIYRQDFRFTLREHSNCSHQNPFLVLTSPSDVKARQAIQVTVTGWGKSWWGYE 99
QY 61 LKTFLLGTTSSAAE---TKEVDOSORHGDIIQKDFLDVYVNLTKTMGIEWVHRFCP 117
Db 100 VLTFFLLGQOAREDTXLSLSEDEHVLGDIIRQDFLDYNNLTUKTMAFRWMEFCP 159
QY 118 QAAFMKTDSDMFINVYDITELLKKNRTRFTFTGFLKNEFFIROPFKSWFVSKSEYWP 177
Db 160 NAKYIMKTDVDFINTGNLVKYLNLNHSKFFTGYPGLIDNYSYRGFYQKTHISYQEYFP 219
QY 178 DRYPPFCSTGYVFSGDVASQVYVSKVPYIKLEDFVGLCLERLNIRLEELHSQPTFF 237
Db 220 KVFPPYCSGLGYIMSRDLVPRYIEMGHVKPIKFEDVYVGCILNLLKVDIHIPEDTNLFF 279
QY 238 PGGLRFSVCLFRRIIVACHFIKPRTLDDYWOAL 269
Db 280 LYRIHLDVCLRRVIAAHGFSKKEIITFWQVM 311
RESULT 10
US-09-739-451-10
; Sequence 10, Application US/09739451
; Patent No. US20010024813A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Dendriac and Brainiac-3
; FILE REFERENCE: PF464
; CURRENT APPLICATION NUMBER: US/09/739,451
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/213,364
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: 60/077,687
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/108,928
; PRIOR FILING DATE: 1998-11-17
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-739-451-10
Query Match 36.6%; Score 548.5; DB 10; Length 320;
Best Local Similarity 38.6%; Pred. No. 5e-47;
Matches 105; Conservative 53; Mismatches 111; Indels 3; Gaps 1;

| | | | |
|---|-----|--|-----|
| Qy | 1 | FKEOSFYKKDGNFLKLPDDCRQTPPFLVLLVTSSHKOLAERMAIRQTWGERVWKGQ | 60 |
| Db | 40 | FYEPEIYRQDFHTUREHNSCHQNPFLVLITSPSDVKARQAIRVTWGEEKSKMWGYE | 99 |
| Qy | 61 | LKTFFLLGTSSAAE---TEVDQESORHGDIIOKDLDVYVNLTKTWGMGLEWHRRFCP | 117 |
| Db | 100 | VLTFFLGOEAEREDKWLALSLEDEHLLYGDIIRDQFELDTYNLNLIKTMAFRWTECP | 159 |
| Qy | 118 | QAAFMKTDSDMFINDVYLTELCLKKNRTTFTFGFLKLINEFPPIOPFSKFVSKSEYPW | 177 |
| Db | 160 | NAKVVMKTDVFNINTGNLVKYLLNLNHSKFTGYPLIDNYSRGFYQKTHISYOEPF | 219 |
| Qy | 178 | DRYPFCSGTGYVSODVASOVNVSKSVPYIKLEDVFGLCLERLINBLELHQPTFF | 237 |
| Db | 220 | KVFPYPCSGLYIMSRDLPRIVEMGHVKPKIFEDVYGVCILNKLKNIHIPEDTNLF | 279 |
| Qy | 238 | PGLRFVSCLFRRIIVACHFKPTRLLDYMQAL | 269 |
| Db | 280 | LYRHLDVCLRRVARAHGFSSKEITTFWQVM | 311 |
| RESULT 11 | | | |
| US-09-992-598-209 | | | |
| ; Sequence 209, Application US/09992598 | | | |
| ; Patent No. US20020160384A1 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: Ashkenazi, Avi J. | | | |
| ; APPLICANT: Baker, Kevin P. | | | |
| ; APPLICANT: Botstein, David | | | |
| ; APPLICANT: Desnoyers, Luc | | | |
| ; APPLICANT: Eaton, Dan L. | | | |
| ; APPLICANT: Ferrara, Napoleone | | | |
| ; APPLICANT: Fong, Sherman | | | |
| ; APPLICANT: Gerber, Hanspeter | | | |
| ; APPLICANT: Gerritsen, Mary E. | | | |
| ; APPLICANT: Goddard, Audrey | | | |
| ; APPLICANT: Godowski, Paul J. | | | |
| ; APPLICANT: Grimaldi, J. Christopher | | | |
| ; APPLICANT: Gurney, Austin L. | | | |
| ; APPLICANT: Kljavin, Ivar J. | | | |
| ; APPLICANT: Napier, Mary A. | | | |
| ; APPLICANT: Pan, James | | | |
| ; APPLICANT: Paoni, Nicholas F. | | | |
| ; APPLICANT: Roy, Margaret Ann | | | |
| ; APPLICANT: Stewart, Timothy A. | | | |
| ; APPLICANT: Tumas, Daniel | | | |
| ; APPLICANT: Watanabe, Colin K. | | | |
| ; APPLICANT: Williams, P. Mickey | | | |
| ; APPLICANT: Wood, William I. | | | |
| ; APPLICANT: Zhang, Zemin | | | |
| ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic | | | |
| ; FILE REFERENCE: P2730PIC20 | | | |
| ; CURRENT APPLICATION NUMBER: US/09/992,598 | | | |
| ; CURRENT FILING DATE: 2001-11-14 | | | |
| ; PRIOR APPLICATION NUMBER: 60/049787 | | | |
| ; PRIOR FILING DATE: 1997-06-16 | | | |
| ; PRIOR APPLICATION NUMBER: 60/062250 | | | |
| ; PRIOR FILING DATE: 1997-10-17 | | | |
| ; PRIOR APPLICATION NUMBER: 60/065186 | | | |
| ; PRIOR FILING DATE: 1997-11-12 | | | |
| ; PRIOR APPLICATION NUMBER: 60/065311 | | | |
| ; PRIOR FILING DATE: 1997-11-13 | | | |
| ; PRIOR APPLICATION NUMBER: 60/066770 | | | |
| ; PRIOR FILING DATE: 1997-11-24 | | | |
| ; PRIOR APPLICATION NUMBER: 60/075945 | | | |
| ; PRIOR FILING DATE: 1998-02-25 | | | |
| ; PRIOR APPLICATION NUMBER: 60/078910 | | | |
| ; PRIOR FILING DATE: 1998-03-20 | | | |
| ; PRIOR APPLICATION NUMBER: 60/083322 | | | |
| ; PRIOR FILING DATE: 1998-04-28 | | | |
| ; PRIOR APPLICATION NUMBER: 60/084600 | | | |
| ; PRIOR FILING DATE: 1998-06-17 | | | |
| ; PRIOR APPLICATION NUMBER: 60/087106 | | | |
| ; PRIOR FILING DATE: 1998-05-28 | | | |
| ; PRIOR APPLICATION NUMBER: 60/087607 | | | |
| ; PRIOR FILING DATE: 1998-06-02 | | | |
| ; PRIOR APPLICATION NUMBER: 60/087609 | | | |
| ; PRIOR FILING DATE: 1998-06-02 | | | |
| ; PRIOR APPLICATION NUMBER: 60/087759 | | | |
| ; PRIOR FILING DATE: 1998-06-02 | | | |
| ; PRIOR APPLICATION NUMBER: 60/087827 | | | |
| ; PRIOR FILING DATE: 1998-06-03 | | | |
| ; PRIOR APPLICATION NUMBER: 60/088021 | | | |
| ; PRIOR FILING DATE: 1998-06-04 | | | |
| ; PRIOR APPLICATION NUMBER: 60/088025 | | | |
| ; PRIOR FILING DATE: 1998-06-04 | | | |
| ; PRIOR APPLICATION NUMBER: 60/088026 | | | |
| ; PRIOR FILING DATE: 1998-06-04 | | | |
| ; PRIOR APPLICATION NUMBER: 60/088028 | | | |
| ; PRIOR FILING DATE: 1998-06-04 | | | |
| ; PRIOR APPLICATION NUMBER: 60/088029 | | | |
| ; PRIOR FILING DATE: 1998-06-04 | | | |
| ; PRIOR APPLICATION NUMBER: 60/088030 | | | |
| ; PRIOR FILING DATE: 1998-06-04 | | | |
| ; PRIOR APPLICATION NUMBER: 60/088033 | | | |
| ; PRIOR FILING DATE: 1998-06-04 | | | |
| ; PRIOR APPLICATION NUMBER: 60/088326 | | | |
| ; PRIOR FILING DATE: 1998-06-04 | | | |
| ; PRIOR APPLICATION NUMBER: 60/088167 | | | |
| ; PRIOR FILING DATE: 1998-06-05 | | | |
| ; PRIOR APPLICATION NUMBER: 60/088202 | | | |
| ; PRIOR FILING DATE: 1998-06-05 | | | |
| ; PRIOR APPLICATION NUMBER: 60/088212 | | | |
| ; PRIOR FILING DATE: 1998-06-05 | | | |
| ; PRIOR APPLICATION NUMBER: 60/088217 | | | |
| ; PRIOR FILING DATE: 1998-06-05 | | | |
| ; PRIOR APPLICATION NUMBER: 60/088655 | | | |
| ; PRIOR FILING DATE: 1998-06-09 | | | |
| ; PRIOR APPLICATION NUMBER: 60/088734 | | | |
| ; PRIOR FILING DATE: 1998-06-10 | | | |
| ; PRIOR APPLICATION NUMBER: 60/088738 | | | |
| ; PRIOR FILING DATE: 1998-06-10 | | | |
| ; PRIOR APPLICATION NUMBER: 60/088742 | | | |
| ; PRIOR FILING DATE: 1998-06-10 | | | |
| ; PRIOR APPLICATION NUMBER: 60/088810 | | | |
| ; PRIOR FILING DATE: 1998-06-10 | | | |
| ; PRIOR APPLICATION NUMBER: 60/088824 | | | |
| ; PRIOR FILING DATE: 1998-06-10 | | | |
| ; PRIOR APPLICATION NUMBER: 60/088826 | | | |
| ; PRIOR FILING DATE: 1998-06-10 | | | |
| ; PRIOR APPLICATION NUMBER: 60/088858 | | | |
| ; PRIOR FILING DATE: 1998-06-11 | | | |
| ; PRIOR APPLICATION NUMBER: 60/088861 | | | |
| ; PRIOR FILING DATE: 1998-06-11 | | | |
| ; PRIOR APPLICATION NUMBER: 60/088876 | | | |
| ; PRIOR FILING DATE: 1998-06-11 | | | |
| ; PRIOR APPLICATION NUMBER: 60/089105 | | | |
| ; PRIOR FILING DATE: 1998-06-12 | | | |
| ; PRIOR APPLICATION NUMBER: 60/089440 | | | |
| ; PRIOR FILING DATE: 1998-06-16 | | | |
| ; PRIOR APPLICATION NUMBER: 60/089512 | | | |
| ; PRIOR FILING DATE: 1998-06-16 | | | |
| ; PRIOR APPLICATION NUMBER: 60/089514 | | | |
| ; PRIOR FILING DATE: 1998-06-16 | | | |
| ; PRIOR APPLICATION NUMBER: 60/089532 | | | |
| ; PRIOR FILING DATE: 1998-06-17 | | | |
| ; PRIOR APPLICATION NUMBER: 60/089538 | | | |
| ; PRIOR FILING DATE: 1998-06-17 | | | |
| ; PRIOR APPLICATION NUMBER: 60/089598 | | | |
| ; PRIOR FILING DATE: 1998-06-17 | | | |
| ; PRIOR APPLICATION NUMBER: 60/089599 | | | |
| ; PRIOR FILING DATE: 1998-06-17 | | | |
| ; PRIOR APPLICATION NUMBER: 60/089600 | | | |
| ; PRIOR FILING DATE: 1998-06-17 | | | |

| | |
|----|-------------------------------------|
| 1 | PRIOR FILING DATE: 1998-06-17 |
| 2 | PRIOR APPLICATION NUMBER: 60/089538 |
| 3 | PRIOR FILING DATE: 1998-06-17 |
| 4 | PRIOR APPLICATION NUMBER: 60/089598 |
| 5 | PRIOR FILING DATE: 1998-06-17 |
| 6 | PRIOR APPLICATION NUMBER: 60/089599 |
| 7 | PRIOR FILING DATE: 1998-06-17 |
| 8 | PRIOR APPLICATION NUMBER: 60/089600 |
| 9 | PRIOR FILING DATE: 1998-06-17 |
| 10 | PRIOR APPLICATION NUMBER: 60/089653 |
| 11 | PRIOR FILING DATE: 1998-06-17 |
| 12 | PRIOR APPLICATION NUMBER: 60/089801 |
| 13 | PRIOR FILING DATE: 1998-06-18 |
| 14 | PRIOR APPLICATION NUMBER: 60/089907 |
| 15 | PRIOR FILING DATE: 1998-06-18 |
| 16 | PRIOR APPLICATION NUMBER: 60/089908 |
| 17 | PRIOR FILING DATE: 1998-06-18 |
| 18 | PRIOR APPLICATION NUMBER: 60/089947 |
| 19 | PRIOR FILING DATE: 1998-06-19 |
| 20 | PRIOR APPLICATION NUMBER: 60/089948 |
| 21 | PRIOR FILING DATE: 1998-06-19 |
| 22 | PRIOR APPLICATION NUMBER: 60/089952 |
| 23 | PRIOR FILING DATE: 1998-06-19 |
| 24 | PRIOR APPLICATION NUMBER: 60/090246 |
| 25 | PRIOR FILING DATE: 1998-06-22 |
| 26 | PRIOR APPLICATION NUMBER: 60/090252 |
| 27 | PRIOR FILING DATE: 1998-06-22 |
| 28 | PRIOR APPLICATION NUMBER: 60/090254 |
| 29 | PRIOR FILING DATE: 1998-06-22 |
| 30 | PRIOR APPLICATION NUMBER: 60/090349 |
| 31 | PRIOR FILING DATE: 1998-06-23 |
| 32 | PRIOR APPLICATION NUMBER: 60/090355 |
| 33 | PRIOR FILING DATE: 1998-06-23 |
| 34 | PRIOR APPLICATION NUMBER: 60/090429 |
| 35 | PRIOR FILING DATE: 1998-06-24 |
| 36 | PRIOR APPLICATION NUMBER: 60/090431 |
| 37 | PRIOR FILING DATE: 1998-06-24 |
| 38 | PRIOR APPLICATION NUMBER: 60/090435 |
| 39 | PRIOR FILING DATE: 1998-06-24 |
| 40 | PRIOR APPLICATION NUMBER: 60/090444 |
| 41 | PRIOR FILING DATE: 1998-06-24 |
| 42 | PRIOR APPLICATION NUMBER: 60/090445 |
| 43 | PRIOR FILING DATE: 1998-06-24 |
| 44 | PRIOR APPLICATION NUMBER: 60/090472 |
| 45 | PRIOR FILING DATE: 1998-06-24 |
| 46 | PRIOR APPLICATION NUMBER: 60/090542 |
| 47 | PRIOR FILING DATE: 1998-06-24 |
| 48 | PRIOR APPLICATION NUMBER: 60/090557 |
| 49 | PRIOR FILING DATE: 1998-06-24 |
| 50 | PRIOR APPLICATION NUMBER: 60/090540 |
| 51 | PRIOR FILING DATE: 1998-06-24 |
| 52 | PRIOR APPLICATION NUMBER: 60/090542 |
| 53 | PRIOR FILING DATE: 1998-06-24 |
| 54 | PRIOR APPLICATION NUMBER: 60/090557 |
| 55 | PRIOR FILING DATE: 1998-06-24 |
| 56 | PRIOR APPLICATION NUMBER: 60/090676 |
| 57 | PRIOR FILING DATE: 1998-06-25 |
| 58 | PRIOR APPLICATION NUMBER: 60/090678 |
| 59 | PRIOR FILING DATE: 1998-06-25 |
| 60 | PRIOR APPLICATION NUMBER: 60/090690 |
| 61 | PRIOR FILING DATE: 1998-06-25 |
| 62 | PRIOR APPLICATION NUMBER: 60/090694 |
| 63 | PRIOR FILING DATE: 1998-06-25 |
| 64 | PRIOR APPLICATION NUMBER: 60/090695 |
| 65 | PRIOR FILING DATE: 1998-06-25 |
| 66 | PRIOR APPLICATION NUMBER: 60/090696 |
| 67 | PRIOR FILING DATE: 1998-06-25 |
| 68 | PRIOR APPLICATION NUMBER: 60/090862 |
| 69 | PRIOR FILING DATE: 1998-06-26 |
| 70 | PRIOR APPLICATION NUMBER: 60/090863 |
| 71 | PRIOR FILING DATE: 1998-06-26 |
| 72 | PRIOR APPLICATION NUMBER: 60/091360 |
| 73 | PRIOR FILING DATE: 1998-07-01 |
| 74 | PRIOR APPLICATION NUMBER: 60/091478 |
| 75 | PRIOR FILING DATE: 1998-07-02 |

;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519

;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 36.6%; Score 548.5; DB 9; Length 331;
Best Local Similarity 38.6%; Pred. No. 5.2e-47;
Matches 105; Conservative 53; Mismatches 111; Indels 3; Gaps 1;

Qy 1 FKEQSFYVKKDGNFLKLPDTCRQTPPFLVLLVTSSHQQLAERMAIROTGWCKERMVKGQ 60
Db 52 FYEYPIYRQDFHFTLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSMMGYE 111
Qy 61 LKTEFLLGTTSSAAE---TKEVDQESQRHGDIIOKDFLDVYYNLTLTMTMGIEWVHRECP 117
Db 112 VLTFLLGQEAKEKEDKMLALSLEDEHLLYGDIIRODFLDYNNLTLTMTMAFRWVTECP 171
Qy 118 QAAFWMTDSDMFINVLYLTETLLKKNRTTRFTGFLKLNFFPIROPFKWFVSKSEYPM 177
Db 172 NAKYVMTDQDVFINVTGNLVKYLNLNHSKFFTGYPIDNYSYRGFTQKTHISYQETPF 231
Qy 178 DRYPPFCSGTGYVFSQDVASQVYVNSKSVPIKLEDFVGLCLERLNIRLEELASQPTFF 237
Db 232 KVFPPYCSGLGYIMSRDLVPRIYEMMGHVKPIKPEDVYVGLCLNLLKVNIIHPIEDTNLFF 291
Qy 238 PGGLRFSVCLFRRIVACHFTKPRTLDDYWOAL 269
Db 292 LYRIHLDVCQLRRVIAAHGFSSKEIITFWQYM 323

Search completed: April 4, 2003, 14:53:31
Job time : 38 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 14:22:18 : Search time 5049.19 Seconds
(without alignments)
15994.713 Million cell updates/sec

Title: US-09-914-152-2

Perfect score: 2775
Sequence: 1 gtgaattctcttctgtc.....aaaaaaaaaaaaaaaaaaaaa 2775

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pi.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pi.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|----------|-------------|
| 1 | 2775 | 100.0 | 2775 | 6 | E38419 | Novel polyp |
| 2 | 2730 | 98.4 | 2762 | 9 | AB020337 | Homo sapi |
| 3 | 2332 | 84.0 | 170121 | 9 | AF064860 | Homo sapi |
| 4 | 2332 | 84.0 | 340000 | 9 | HS21C080 | Homo sapi |
| 5 | 2329.4 | 83.9 | 10562 | 6 | E38420 | Novel polyp |
| 6 | 933 | 33.6 | 933 | 9 | AF145784 | Homo sapi |
| 7 | 924 | 33.3 | 933 | 9 | HS46078 | Homo sapi |
| 8 | 899.2 | 32.4 | 1576 | 9 | AB041416 | Homo sapi |
| 9 | 882.6 | 31.8 | 1570 | 9 | AB041415 | Pan panis |
| 10 | 879.6 | 31.7 | 1565 | 9 | AB041413 | Homo sapi |
| 11 | 872.4 | 31.4 | 1360 | 9 | AB041412 | Gorilla g |
| 12 | 869.6 | 31.3 | 1566 | 9 | AB041414 | Pan trogl |
| 13 | 850.4 | 30.6 | 1579 | 9 | AB041417 | Pongo pyg |
| 14 | 542.2 | 19.5 | 196900 | 2 | AC020851 | Mus muscu |
| 15 | 534.4 | 19.3 | 175861 | 2 | AC120346 | Mus muscu |
| 16 | 520.8 | 18.8 | 927 | 10 | AF254738 | Mus muscu |
| 17 | 341.8 | 12.3 | 65400 | 2 | AC109264 | Mus muscu |
| 18 | 185.2 | 6.7 | 917 | 10 | AB039137 | Mus muscu |
| 19 | 185.2 | 6.7 | 917 | 10 | AB039140 | Mus muscu |
| 20 | 185.2 | 6.7 | 917 | 10 | AB039141 | Mus muscu |
| 21 | 185.2 | 6.7 | 917 | 10 | AB039142 | Mus muscu |
| 22 | 183.6 | 6.6 | 917 | 10 | AB039136 | Mus muscu |
| 23 | 183.6 | 6.6 | 917 | 10 | AB039138 | Mus muscu |
| 24 | 183.6 | 6.6 | 917 | 10 | AB039143 | Mus spici |
| 25 | 182 | 6.6 | 917 | 10 | AB039134 | Mus muscu |
| 26 | 182 | 6.6 | 917 | 10 | AB039135 | Mus muscu |
| 27 | 182 | 6.6 | 917 | 10 | AB039139 | Mus muscu |
| 28 | 182 | 6.6 | 1175 | 10 | AF029790 | Mus muscu |
| 29 | 182 | 6.6 | 219368 | 10 | AC098721 | Mus muscu |
| 30 | 166 | 6.0 | 1053 | 9 | AB041409 | Pan panis |
| 31 | 165 | 5.9 | 1056 | 9 | AB041410 | Gorilla g |
| 32 | 165 | 5.9 | 1059 | 9 | AB041411 | Pongo pyg |
| 33 | 164.4 | 5.9 | 1060 | 9 | AB041408 | Pan trogl |
| 34 | 163.4 | 5.9 | 1037 | 9 | AB041407 | Homo sapi |
| 35 | 163.4 | 5.9 | 1739 | 6 | E07739 | cDNA encodi |
| 36 | 163.4 | 5.9 | 202001 | 9 | AC016723 | Homo sapi |
| 37 | 161.8 | 5.8 | 1909 | 9 | AF117222 | Homo sapi |
| 38 | 151.8 | 5.5 | 137703 | 9 | AC092036 | Homo sapi |
| 39 | 151.8 | 5.5 | 159095 | 2 | AC016935 | Homo sapi |
| 40 | 148.8 | 5.4 | 161824 | 9 | AC096721 | Homo sapi |
| 41 | 142.8 | 5.1 | 996 | 10 | AB039156 | Mus muscu |
| 42 | 142.8 | 5.1 | 996 | 10 | AB039158 | Mus muscu |
| 43 | 142.8 | 5.1 | 996 | 10 | AB039162 | Mus muscu |
| 44 | 142.8 | 5.1 | 996 | 10 | AB039163 | Mus spici |
| 45 | 142 | 5.1 | 164076 | 9 | AC008006 | Homo sapi |

ALIGNMENTS

RESULT 1
E38419
LOCUS E38419 Novel polypeptide.
DEFINITION E38419
ACCESSION E38419
VERSION E38419.1 GI:18626993
KEYWORDS JP 2000245464-A/1.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2775)
AUTHORS Narimatsu, H., Isshiki, S., Togayachi, A. and Sasaki, K.
TITLE Novel polypeptide
JOURNAL Patent: JP 2000245464-A 1 12-SEP-2000;

2775 bp DNA linear PAT 31-JAN-2002

Mon Apr 14 10:18:02 2003

| | |
|------------|--|
| COMMENT | KYOWA HAKKO KOGYO CO LTD OS Homo sapiens (human) PN JP 2000245464 -A/1 PD 12-SEP-2000 PF 25-FEB-1999 JP 19990047571 PR PI HISASHI NARIMATSU, SOICHIRO ISSHIKI, AKIRA TOGAYAUCHI, PI KATSUTOSHI SASAKI PC C12N15/09, A01K67/027, C12N1/21, C12N5/10, C12N9/10, C12P19/00, PC C12P21/02, C12Q1/68, C01N33/53/(C12N1/21, C12R1:185), (C12N5/10, PC C12P21/08, C12R1:91), PC C12R1:91), PC (C12P21/02, C12R1:185), (C12P21/02, C12R1:91), C12N15/00, C12N5/00, PC (C12N5/00, C12R1:91) CC FH Key Location/Qualifiers FT CDS (402), (1331). |
| FEATURES | source 1. .2775 /organism="Homo sapiens" /db_xref="taxon:9606" 681 a 698 c 669 g 727 t |
| BASE COUNT | 681 a 698 c 669 g 727 t |
| ORIGIN | Query Match 100.0%; Score 2775; DB 6; Length 2775; Best Local Similarity 100.0%; Pred. No. 0; Matches 2775; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |
| QY | 1 GTGAATTCCTCTCTCTGCTGAGCTGGGATATTCTTCTCTCTGCTGCTTGGACATCA 60 |
| Db | 1 GTGAATTCCTCTCTCTGCTGAGCTGGGATATTCTTCTCTCTGCTGCTTGGACATCA 60 |
| QY | 61 GAGCTGAGGCTCTCTGGCTTTGGACCCGAGGATTTATACCAAGCAGGTTTCTGGGTTTC 120 |
| Db | 61 GAGCTGAGGCTCTCTGGCTTTGGACCCGAGGATTTATACCAAGCAGGTTTCTGGGTTTC 120 |
| QY | 121 TCAGGCTTTGGCTTTGGCTGATAGTTACACCATTTGGCATATCTGGTTCTTGAGGCTCTT 180 |
| Db | 121 TCAGGCTTTGGCTTTGGCTGATAGTTACACCATTTGGCATATCTGGTTCTTGAGGCTCTT 180 |
| QY | 181 GGTCTTGACTGAGCCACACTCTCTGGCATCCAGCGTCTCCAGTTGCATGGCTGTAC 240 |
| Db | 181 GGTCTTGACTGAGCCACACTCTCTGGCATCCAGCGTCTCCAGTTGCATGGCTGTAC 240 |
| QY | 241 GTGATTCCTGTGAGATCACCATTCTTTGGTAAACAAACCAAGCCAGAACCTGATAATTA 300 |
| Db | 241 GTGATTCCTGTGAGATCACCATTCTTTGGTAAACAAACCAAGCCAGAACCTGATAATTA 300 |
| QY | 301 TGGAGCATCTACACTGACAGTTCTTTGAGACAAATTTCTCTTTGGCATTTACACTGTGG 360 |
| Db | 301 TGGAGCATCTACACTGACAGTTCTTTGAGACAAATTTCTCTTTGGCATTTACACTGTGG 360 |
| QY | 361 CTTTAGCTTTCAAAACAGAGGTTCTCTTACCAGCAAAATGGCTTTCCGGAAGATGA 420 |
| Db | 361 CTTTAGCTTTCAAAACAGAGGTTCTCTTACCAGCAAAATGGCTTTCCGGAAGATGA 420 |
| QY | 421 GATTGATGATATTTGGCTTCTGTTCTGGGGCTCTTTGTTGTTATTTAGCATGTACA 480 |
| Db | 421 GATTGATGATATTTGGCTTCTGTTCTGGGGCTCTTTGTTGTTATTTAGCATGTACA 480 |
| QY | 481 GTCTAAATCTTTCAAGAACAGTCTTTGTTTACAAGAAACGGGAACTTCTTAAGC 540 |
| Db | 481 GTCTAAATCTTTCAAGAACAGTCTTTGTTTACAAGAAACGGGAACTTCTTAAGC 540 |
| QY | 541 TCCAGATACAGACTGACGGCAGACACTCTCTCTCTCTGCTGCTGGTGAACCTCATCC 600 |
| Db | 541 TCCAGATACAGACTGACGGCAGACACTCTCTCTCTCTGCTGCTGGTGAACCTCATCC 600 |
| QY | 601 ACAAAACAGTTGGCTGAGCGCATCGCCATCCGGCAGACGTTGGGGAAAGAGAGATGGTGA 660 |
| Db | 601 ACAAAACAGTTGGCTGAGCGCATCGCCATCCGGCAGACGTTGGGGAAAGAGAGATGGTGA 660 |
| QY | 661 AGGAAAGCAGCTGAAGACATTTCTCTCTCTGGGGACCAACCAAGCAGTGCAGCGGAACGA 720 |

| | | | |
|----|------|--|------|
| b | 1741 | CCTGACTTGAACCACTCCACAGTCTGCCCTCCCTTAGGAGGGACACGTGATGATGAGGTC | 1800 |
| | | 1742 | 1800 |
| y | 1801 | TCGGAGCCGGCATCCTCCATCCCTGTGAGTCCCTCCACCTCAGCTCCAGTCCCTTGT | 1860 |
| | | 1802 | 1860 |
| ib | 1801 | TCGGAGCCGGCATCCTCCATCCCTGTGAGTCCCTCCACCTCAGCTCCAGTCCCTTGT | 1860 |
| | | 1802 | 1860 |
| y | 1861 | GCCTTTTGGAGCTAAGCTTGGGATGACCAAAATCACCCAGCTCCTTCATTTCACAGGGCT | 1920 |
| | | 1862 | 1920 |
| Db | 1861 | GCCTTTTGGAGCTAAGCTTGGGATGACCAAAATCACCCAGCTCCTTCATTTCACAGGGCT | 1920 |
| | | 1862 | 1920 |
| Qy | 1921 | GGATGTAGCTTGGGATTTAGTTCATGTTATCGGCTCGGTACTCAACACAAACCAAGTTTCA | 1980 |
| | | 1922 | 1980 |
| Db | 1921 | GGATGTAGCTTGGGATTTAGTTCATGTTATCGGCTCGGTACTCAACACAAACCAAGTTTCA | 1980 |
| | | 1922 | 1980 |
| Qy | 1981 | TCCGAGGAAATGTCGCCGAGTGGATCGAGCTCACATGCTGAGGAACACCCAGCTCTGGA | 2040 |
| | | 1982 | 2040 |
| Db | 1981 | TCCGAGGAAATGTCGCCGAGTGGATCGAGCTCACATGCTGAGGAACACCCAGCTCTGGA | 2040 |
| | | 1982 | 2040 |
| Qy | 2041 | CAGAGTTCTTATAAATGTATAAATAGGCTCAGAAACCACTGCAATTCCTGACCTGCTGTAC | 2100 |
| | | 2042 | 2100 |
| Db | 2041 | CAGAGTTCTTATAAATGTATAAATAGGCTCAGAAACCACTGCAATTCCTGACCTGCTGTAC | 2100 |
| | | 2042 | 2100 |
| Qy | 2101 | AGACTGCCACACACTGCTGACCTGCTAGCGAGGACATCCCTTCTGAGCCATCTGCTG | 2160 |
| | | 2102 | 2160 |
| Db | 2101 | AGACTGCCACACACTGCTGACCTGCTAGCGAGGACATCCCTTCTGAGCCATCTGCTG | 2160 |
| | | 2102 | 2160 |
| Qy | 2161 | CTCTCTCATTTGATCACCCTCACTGCTCCCTGTTTGTATCAATGGGACCAACCACTGC | 2220 |
| | | 2162 | 2220 |
| Db | 2161 | CTCTCTCATTTGATCACCCTCACTGCTCCCTGTTTGTATCAATGGGACCAACCACTGC | 2220 |
| | | 2162 | 2220 |
| Qy | 2221 | CCCAGGACACTTTAGGGCTCTCAGTTCAAATGAAGGACAGTTGAATCAGATGGGTT | 2280 |
| | | 2222 | 2280 |
| Db | 2221 | CCCAGGACACTTTAGGGCTCTCAGTTCAAATGAAGGACAGTTGAATCAGATGGGTT | 2280 |
| | | 2222 | 2280 |
| Qy | 2281 | CATGTGGGATTTGGGACCTTCTGGGAAATTCAGTTGGAGTCAAGTCAGGATGCTCTCAA | 2340 |
| | | 2282 | 2340 |
| Db | 2281 | CATGTGGGATTTGGGACCTTCTGGGAAATTCAGTTGGAGTCAAGTCAGGATGCTCTCAA | 2340 |
| | | 2282 | 2340 |
| Qy | 2341 | GGACCCCTCGGCTCAGAGCCCTAAAGTGGGCCCTGGTGAAGCAGGGTGGTCCGCTCC | 2400 |
| | | 2342 | 2400 |
| Db | 2341 | GGACCCCTCGGCTCAGAGCCCTAAAGTGGGCCCTGGTGAAGCAGGGTGGTCCGCTCC | 2400 |
| | | 2342 | 2400 |
| Qy | 2401 | ACTTCCCAAGCCTGAGCCAGCTCATCTTCATTGAATGTCTCATTTGGCCGAGGAACAAC | 2460 |
| | | 2402 | 2460 |
| Db | 2401 | ACTTCCCAAGCCTGAGCCAGCTCATCTTCATTGAATGTCTCATTTGGCCGAGGAACAAC | 2460 |
| | | 2402 | 2460 |
| Qy | 2461 | TGAACCTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT | 2520 |
| | | 2462 | 2520 |
| Db | 2461 | TGAACCTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT | 2520 |
| | | 2462 | 2520 |
| Qy | 2521 | TCGTGGGAGCCTGTGTGGAGGGTGTATATAACCAAGGACTTCGTTAGTTTTCGCCATT | 2580 |
| | | 2522 | 2580 |
| Db | 2521 | TCGTGGGAGCCTGTGTGGAGGGTGTATATAACCAAGGACTTCGTTAGTTTTCGCCATT | 2580 |
| | | 2522 | 2580 |
| Qy | 2581 | CAGCCATGTCAGCTGACATGCAAAAGTAACTTGTCTTAATATAGAAATGATTTTCT | 2640 |
| | | 2582 | 2640 |
| Db | 2581 | CAGCCATGTCAGCTGACATGCAAAAGTAACTTGTCTTAATATAGAAATGATTTTCT | 2640 |
| | | 2582 | 2640 |
| Qy | 2641 | TTTAAATTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAA | 2700 |
| | | 2642 | 2700 |
| Db | 2641 | TTTAAATTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAA | 2700 |
| | | 2642 | 2700 |
| Qy | 2701 | TGTCACATATAAATGTTGGACTAACTCTTAAAAAATAAATAAATAAATAAATAAATAAATAA | 2760 |
| | | 2702 | 2760 |
| Db | 2701 | TGTCACATATAAATGTTGGACTAACTCTTAAAAAATAAATAAATAAATAAATAAATAAATAA | 2760 |
| | | 2702 | 2760 |
| Qy | 2761 | AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 2775 |
| | | 2762 | 2775 |
| Db | 2761 | AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 2775 |
| | | 2762 | 2775 |

1120 AACTGGAAGACGTGTTGTGGGCTCTGCTCGAAGGCTGAACATCAGATTGAGGAGC 1179
|||||
90766 AACTGGAAGACGTGTTGTGGGCTCTGCTCGAAGGCTGAACATCAGATTGAGGAGC 90825
|||||
1180 TCACACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCCTCCGATGCCCTTTCAGGA 1239
|||||
90826 TCACACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCCTCCGATGCCCTTTCAGGA 90885
|||||
1240 GGATCGTGGCCCTGCCACTTCATCAAGCTCGGACTCTCTTGAGCTACTGGCAGGCTCTAG 1299
|||||
90886 GGATCGTGGCCCTGCCACTTCATCAAGCTCGGACTCTCTTGAGCTACTGGCAGGCTCTAG 90945
|||||
1300 AGAATTCGGGGGGAAGATGTCGCCCTGTCTGAGGGGAGCCAGAGGCACATCGGGAC 1359
|||||
90946 AGAATTCGGGGGGAAGATGTCGCCCTGTCTGAGGGGAGCCAGAGGCACATCGGGAC 91005
|||||
1360 AAGTTTTCAGATAAACCCTGGGGATAGTTTTCCTAGATTTTGAAGAGGGGCGGGACAG 1419
|||||
91006 AAGTTTTCAGATAAACCCTGGGGATAGTTTTCCTAGATTTTGAAGAGGGGCGGGACAG 91065
|||||
1420 AGGATGCTGTTCTTCAGTCTGCTGAATCCACGCCCAATATAGCCCGTCTCTTGGGCAGCACACTCT 1539
|||||
91066 AGGATGCTGTTCTTCAGTCTGCTGAATCCACGCCCAATATAGCCCGTCTCTTGGGCAGCACACTCT 91185
|||||
1480 ATTAGTTCCTACTGGTGGTCCCGCAGCAATATAGCCCGTCTCTTGGGCAGCACACTCT 1539
|||||
91126 ATTAGTTCCTACTGGTGGTCCCGCAGCAATATAGCCCGTCTCTTGGGCAGCACACTCT 91185
|||||
1540 TCATACTAAGTGTGTGACATACACCTGGATTTTTCAGGGGTGAGTATCCTATGA 1599
|||||
91186 TCATACTAAGTGTGTGACATACACCTGGATTTTTCAGGGGTGAGTATCCTATGA 91245
|||||
1600 CATGATGGGTGTTACCATCTCTAAATTTACAGGCAAGGACAGCAGCTGCAGAGGTACA 1659
|||||
91246 CATGATGGGTGTTACCATCTCTAAATTTACAGGCAAGGACAGCAGCTGCAGAGGTACA 91305
|||||
1660 GAAACTTGTCCCAAGGCTCAGCCAGTATAGGAGGGGAATGAAATTCAGGACT 1719
|||||
91306 GAAACTTGTCCCAAGGCTCAGCCAGTATAGGAGGGGAATGAAATTCAGGACT 91365
|||||
1720 GTCAGAAATCTGGTGGGAGCCCTGACTTGAACCACTCCACAGTGTGCTCCCTTAGGA 1779
|||||
91366 GTCAGAAATCTGGTGGGAGCCCTGACTTGAACCACTCCACAGTGTGCTCCCTTAGGA 91425
|||||
1780 GGGGACACTGATGATGAGTCTCGGAGCCGGCATCTTCCATCCCTGTCAGTCCCTCC 1839
|||||
91426 GGGGACACTGATGATGAGTCTCGGAGCCGGCATCTTCCATCCCTGTCAGTCCCTCC 91485
|||||
1840 ACCTCAGCTCCCAAGTCTGCTGCTTTTGGAGCTAAGCTGGGATGACCAAAATTCACCC 1899
|||||
91486 ACCTCAGCTCCCAAGTCTGCTGCTTTTGGAGCTAAGCTGGGATGACCAAAATTCACCC 91545
|||||
1900 AGCTCCTTCATTACAGGGCTGGATGATGCTGGGATTTAGTCCATCTTATCGGCTCGGTA 1959
|||||
91546 AGCTCCTTCATTACAGGGCTGGATGATGCTGGGATTTAGTCCATCTTATCGGCTCGGTA 91605
|||||
1960 CTCACACACCAAGTTTCATCCGAGGAATGTCCCGCAGTGGATGATGATGATGATGATG 2019
|||||
91606 CTCACACACCAAGTTTCATCCGAGGAATGTCCCGCAGTGGATGATGATGATGATGATG 91665
|||||
2020 TGAGGAACACCCAGCTCTGGACAGAGTCTTATAATGTATAATGATTAATAGCTCAGAAACA 2079
|||||
91666 TGAGGAACACCCAGCTCTGGACAGAGTCTTATAATGTATAATGATTAATAGCTCAGAAACA 91725
|||||
2080 CTGCAATCTGACCTGCTGTACAGAGTCCGACACTGCTGACCTGCTGACCTGCTGACCTG 2139
|||||
91726 CTGCAATCTGACCTGCTGTACAGAGTCCGACACTGCTGACCTGCTGACCTGCTGACCTG 91785
|||||
2140 TCCCTTCTGAGCCATCTGCTGCTCTCTCATTTTCATCACCCCACTGCTCCCTGTTTGA 2199
|||||
91786 TCCCTTCTGAGCCATCTGCTGCTCTCTCATTTTCATCACCCCACTGCTCCCTGTTTGA 91845
|||||

QY 2200 TCAATGGGGACCAAGCCACTGCCCCAGGAGCACTTTAGGCTCTCAGTTCAAACTGAAGGA 2259
|||||
Db 91846 TCAATGGGGACCAAGCCACTGCCCCAGGAGCACTTTAGGCTCTCAGTTCAAACTGAAGGA 91905
|||||
QY 2260 CACTTGAACCTCAGATGGGGTTCATGTGGGATCTGGGAGCTTTCTGGGAATTCAGTTGGA 2319
|||||
Db 91906 CAGTTGAACCTCAGATGGGGTTCATGTGGGATCTGGGAGCTTTCTGGGAATTCAGTTGGA 91965
|||||
QY 2320 GTCAGCTCAGGATGCTCTCAAGGACCCCTCGGGCTCAGAGCCCTAAAGTGGGCCCTGGTG 2379
|||||
Db 91966 GTCAGCTCAGGATGCTCTCAAGGACCCCTCGGGCTCAGAGCCCTAAAGTGGGCCCTGGTG 92025
|||||
QY 2380 AAGCAGGGTGGTCTCGCTCCACTTCCCAAGCTGAGCCCAAGCTCATCTTCATTGAATGT 2439
|||||
Db 92026 AAGCAGGGTGGTCTCGCTCCACTTCCCAAGCTGAGCCCAAGCTCATCTTCATTGAATGT 92085
|||||
QY 2440 CTCATTGGCCCGAGGAGCAACTGAACTTTGTGGTGTGCTTTAGCCTTCAGTTGCTCC 2499
|||||
Db 92086 CTCATTGGCCCGAGGAGCAACTGAACTTTGTGGTGTGCTTTAGCCTTCAGTTGCTCC 92145
|||||
QY 2500 GCTGCTCTCCTACCCAGAGGTTTGTGCGAGCCCTGTGTGAGGGTGTGTATAAACCAAGGT 2559
|||||
Db 92146 GCTGCTCTCCTACCCAGAGGTTTGTGCGAGCCCTGTGTGAGGGTGTGTATAAACCAAGGT 92205
|||||
QY 2560 ACTTCGTTAGTTTTCCTCCACTCAGCCATGGTCACGTGACATGCAAAAGTAATCTGCTCT 2619
|||||
Db 92206 ACTTCGTTAGTTTTCCTCCACTCAGCCATGGTCACGTGACATGCAAAAGTAATCTGCTCT 92265
|||||
QY 2620 AATTATAGAAATGATTTTCTTTAAATTTTACTTTACCAGACTTTACTTTGTACTCAG 2679
|||||
Db 92266 AATTATAGAAATGATTTTCTTTAAATTTTACTTTACCAGACTTTACTTTGTACTCAG 92325
|||||
QY 2680 AGAAGAGCCCTCAGATGCTGTGCACATATAATGTTGGAGCTAAACTCTTAAAAAAA 2739
|||||
Db 92326 AGAAGAGCCCTCAGATGCTGTGCACATATAATGTTGGAGCTAAACTCTTAACTTGA 92385
|||||
QY 2740 AAAAAAA 2747
|||||
Db 92386 CAGCAAAA 92393

RESULT 4
HS21C080
LOCUS
DEFINITION Homo sapiens chromosome 21 segment HS21C080.
ACCESSION AL163280 AP001735 BA000005
VERSION AL163280.2 GI:7717369

KEYWORDS

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 340000)

REFERENCE

AUTHORS

Hattori, M., Fujiyama, A., Taylor, T. D., Watanabe, H., Yada, T.,
Park, H. S., Toyoda, A., Ishii, K., Totoki, Y., Choi, D. K., Soeda, E.,
Ohki, M., Takagi, T., Sakaki, Y., Taudien, S., Blechschmidt, K.,
Polly, A., Menzel, U., Delabar, J., Kumpf, K., Lehmann, R.,
Patterson, D., Reichwald, K., Rump, A., Schillhabel, M., Schudy, A.,
Zimmermann, W., Rosenthal, A., Kudoh, J., Shibuya, K., Kawasaki, K.,
Asakawa, S., Shintani, A., Sasaki, T., Nagamine, K., Mitsuyama, S.,
Antonarakis, S. E., Minoshima, S., Shimizu, N., Nordstie, G.,
Hornischer, K., Brandt, P., Scharfe, M., Schoen, O., Desario, A.,
Reichelt, J., Kauer, G., Bloeker, H., Ramser, J., Beck, A., Klages, S.,
Hennig, S., Riesselmann, L., Dagand, E., Wehrmeyer, S., Borzym, K.,
Gardiner, K., Nizetic, D., Francis, F., Lehnach, H., Reinhardt, R., and
Yaspo, M. L.

Direct Submission

TITLE

JOURNAL

Submitted (05-MAY-2000) The Chromosome 21 Mapping and Sequencing
Consortium: * RIKEN Genomic Sciences Center, Human Genome Research
Group * Institute of Molecular Biotechnology, Genome Analysis *
Keio University School of Medicine, Dept. of Molecular Biology *
GBF, Dept. of Genome Analysis * Max-Planck Institute for Molecular
Genetics (addresses see below)

COMMENT

The Chromosome 21 Mapping and Sequencing Consortium consists of
 * RIKEN Genomic Sciences Center, Human Genome Research Group, *
 Sagamihara 228-8555, Japan,
 * e-mail: sakaki@gsc.riken.go.jp
 * URL: http://hgp.gsc.riken.go.jp/
 and
 * Institute of Molecular Biotechnology, Genome Analysis, *
 Beutenbergstrasse 11, D-07745 Jena, Germany,
 * e-mail: gscj-submit@genome.imb-jena.de
 * URL: http://genome.imb-jena.de/
 and
 * Keio University School of Medicine, Dept. of Molecular Biology, *
 Tokyo 160-8582, Japan,
 * e-mail: shimizu@db.med.keio.ac.jp
 * URL: http://adenine.dmb.med.keio.ac.jp/
 and
 * GBF, Dept. of Genome Analysis,
 * Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e-mail:
 info.genome@gbf.de
 * URL: http://genome.gbf.de/
 and
 * Max-Planck Institute for Molecular Genetics,
 * Ihnestrasse 73, D-14195 Berlin, Germany,
 * e-mail: info-chr21@molgen.mpg.de
 * URL: http://chr21.rz-berlin.mpg.de/.

FEATURES

source

1. 340000
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.3"
 <1. 125946

source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.3"
 /clone="P160L9, 5' partial"
 /clone_lib="RPC11,3-5 PAC library"
 /note="Accession No. AF121897"
 73410. 243533
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.3"
 /clone="P70124"
 /clone_lib="RPC11,3-5 PAC library"
 /note="Accession No. AF064860"
 224137. >340000
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.3"

source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.3"
 /clone="P206A10, 3' partial"
 /clone_lib="RPC11,3-5 PAC library"
 /note="Accession No. AF121782"
 286628. >340000
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.3"
 /clone="BAC-291B3, 3' partial"
 /clone_lib="BAC library"
 /note="Accession No. AF064857"
 1088. 1566
 /note="L1M4"

source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.3"
 /clone="BAC-291B3, 3' partial"
 /clone_lib="BAC library"
 /note="Accession No. AF064857"
 1088. 1566
 /note="L1M4"

repeat_region

/rpt_family="LINE/L1"
 /rpt_type=DISPERSED
 2398. 2495

repeat_region

/note="(TTTA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 complement(2496. 2878)

repeat_region

/note="THE1C"
 complement(2496. 2878)

repeat_region
 /rpt_family="L1/L2"
 /rpt_type=DISPERSED
 2879. 2931

repeat_region
 /note="(TCTA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 2932. 3004

repeat_region
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3005. 3074

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

gene
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157

```
/rpt_family="SINE/Alu"  
/rpt_type=DISPERSED  
complement(11315..11556)  
/note="AluJo"  
/rpt_family="SINE/Alu"  
/rpt_type=DISPERSED  
11702..11753  
/note="MIR"  
/rpt_family="SINE/MIR"  
/rpt_type=DISPERSED  
13649..12678  
/gene="SH3BGR"  
/number=5  
complement(13115..13425)  
/note="AluSx"  
/rpt_family="SINE/Alu"  
/rpt_type=DISPERSED  
complement(13598..13892)  
/note="AluY"  
/rpt_family="SINE/Alu"  
/rpt_type=DISPERSED  
14013..14131  
/note="(CA)n"  
/rpt_family="Simple_repeat"  
  
Query Match 84.0%; Score 2332; DB 9; Length 340000;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 2338; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 400 AATGGCTTTCCGAAGATGAGATTGATGTATATTTGCCCTTCGTCTGGGGCTCTTT 459  
Db 163454 AGATGGCTTTCCGAAGATGAGATTGATGTATATTTGCCCTTCGTCTGGGGCTCTTT 163513  
  
QY 460 GTTTGTATTTAGCATGTACAGTCTAAATCCTTTCAAGAAACAGTCCCTTTGTATTACAGA 519  
Db 163514 GTTTGTATTTAGCATGTACAGTCTAAATCCTTTCAAGAAACAGTCCCTTTGTATTACAGA 163573  
  
QY 520 AAGACGGGAATTCCTTAAGTCCAGATACAGACTGCAGCGAGACACTCCCTTCCTCG 579  
Db 163574 AAGACGGGAATTCCTTAAGTCCAGATACAGACTGCAGCGAGACACTCCCTTCCTCG 163633  
  
QY 580 TCCTGCTGTGACCTCATCCACAAACAGTTGGCTGAGCGCATGGCCATCGGCGAGACGT 639  
Db 163634 TCCTGCTGTGACCTCATCCACAAACAGTTGGCTGAGCGCATGGCCATCGGCGAGACGT 163693  
  
QY 640 GGGGAAAGAGAGATGGTGAAGGAAAGCAGCTGAAGACATTTCTCCTCCTGGGACCA 699  
Db 163694 GGGGAAAGAGAGATGGTGAAGGAAAGCAGCTGAAGACATTTCTCCTCCTGGGACCA 163753  
  
QY 700 CCAGCAGTGCAGCGGAAACGAAAGAGTGGACAGGAGAGCCAGCGACACGGGAGACATTA 759  
Db 163754 CCAGCAGTGCAGCGGAAACGAAAGAGTGGACAGGAGAGCCAGCGACACGGGAGACATTA 163813  
  
QY 760 TCAGAAAGATTTCCTAGAGCTTATTACAACTGACCCCTGAAGACCATGATGGGCATAG 819  
Db 163814 TCAGAAAGATTTCCTAGAGCTTATTACAACTGACCCCTGAAGACCATGATGGGCATAG 163873  
  
QY 820 AATGGGTCCATCGCTTTGCTCCTCAGCGCGCGTTTGTGTGATGAAGAAACAGACTCAGACATGT 879  
Db 163874 AATGGGTCCATCGCTTTGCTCCTCAGCGCGCGTTTGTGTGATGAAGAAACAGACTCAGACATGT 163933  
  
QY 880 TCATCAATGTGTACTCTGACTGAACTGCTTCTGAAGAAAAACAGAAACACCAGCTTTT 939  
Db 163934 TCATCAATGTGTACTCTGACTGAACTGCTTCTGAAGAAAAACAGAAACACCAGCTTTT 163993  
  
QY 940 TCAGTGGCTTCCTGAAACTCAATGAGTTTCCCATCAGGAGGCCATTTCAGCAAGTGGTTTG 999  
Db 163994 TCAGTGGCTTCCTGAAACTCAATGAGTTTCCCATCAGGAGGCCATTTCAGCAAGTGGTTTG 164053  
  
QY 1000 TCAGTAAATCTGAATCCGTGGACAGTACCACCATTTCTGCTCCGGACCGGGCTACG 1059  
Db 164054 TCAGTAAATCTGAATCCGTGGACAGTACCACCATTTCTGCTCCGGACCGGGCTACG 164113
```

```
QY 1060 TGTTCCTGGCGACGTGGCGAGTCAAGTGTACAAATGTCTCCAAGAGCGTCCCATACATTA 1119  
Db 164114 TGTTCCTGGCGACGTGGCGAGTCAAGTGTACAAATGTCTCCAAGAGCGTCCCATACATTA 164173  
  
QY 1120 AACTGGAAGACGTGTTTGTGGGGCTCTGCCCTCGAAAGGCTCAACATCAGATTGGAGGAGC 1179  
Db 164174 AACTGGAAGACGTGTTTGTGGGGCTCTGCCCTCGAAAGGCTCAACATCAGATTGGAGGAGC 164233  
  
QY 1180 TCCACTCCACGACCGACCTTTTCCAGGGGCTTACGCTTCTCCGTATGCCTCTTCACAGA 1239  
Db 164234 TCCACTCCACGACCGACCTTTTCCAGGGGCTTACGCTTCTCCGTATGCCTCTTCACAGA 164293  
  
QY 1240 GGATCGTGGCTGGCACTTCATCAAGCTTCGAGCTCTCTTGGAGTACTTGGCAGGCTCTAG 1299  
Db 164294 GGATCGTGGCTGGCACTTCATCAAGCTTCGAGCTCTCTTGGAGTACTTGGCAGGCTCTAG 164353  
  
QY 1300 AGAATTCGCGGGGGAAGATTGTCCGCTGTCTGAGGGGAGCCAGAGGCAACATCCGGAC 1359  
Db 164354 AGAATTCGCGGGGGAAGATTGTCCGCTGTCTGAGGGGAGCCAGAGGCAACATCCGGAC 164413  
  
QY 1360 AAGTTTCAGATAACCCGTTGGGATAGTTTTCCTAGATTTCGAAAGAGGGGGGAGACAG 1419  
Db 164414 AAGTTTCAGATAACCCGTTGGGATAGTTTTCCTAGATTTCGAAAGAGGGGGGAGACAG 164473  
  
QY 1420 AGGATGCTGTTCTTCAAGTGTGCTGAAATCCACGCCAGAAATGTCCGTTTCATGAAGTCACTG 1479  
Db 164474 AGGATGCTGTTCTTCAAGTGTGCTGAAATCCACGCCAGAAATGTCCGTTTCATGAAGTCACTG 164533  
  
QY 1480 ATTAGTTCCCACTTGGTGGCCCCAGGCAATTAATAGCCCGTCTCTTGGGCACGCACTCT 1539  
Db 164534 ATTAGTTCCCACTTGGTGGCCCCAGGCAATTAATAGCCCGTCTCTTGGGCACGCACTCT 164593  
  
QY 1540 TCATACTAAGTGTGTGACATACACTGATTTTTCGATTTTCAGGGTTCAGTATCCTATGA 1599  
Db 164594 TCATACTAAGTGTGTGACATACACTGATTTTTCGATTTTCAGGGTTCAGTATCCTATGA 164653  
  
QY 1600 CATGATGGGTGTTACCATCTCTAATTTTACAGGCAAGGACACAGCAGCTGCGAGAGGTACA 1659  
Db 164654 CATGATGGGTGTTACCATCTCTAATTTTACAGGCAAGGACACAGCAGCTGCGAGAGGTACA 164713  
  
QY 1660 GAAACTTGTCCCAAGGCTCACAGCCAGTAGGCATAGAGCGGGAATGAAATTCAGCACT 1719  
Db 164714 GAAACTTGTCCCAAGGCTCACAGCCAGTAGGCATAGAGCGGGAATGAAATTCAGCACT 164773  
  
QY 1720 GTCAGAACTCTGGTGGGAGCCCTGACTTGAACCACTCCACAGTCTGCCCTCCCTTAGGA 1779  
Db 164774 GTCAGAACTCTGGTGGGAGCCCTGACTTGAACCACTCCACAGTCTGCCCTCCCTTAGGA 164833  
  
QY 1780 GGGGACACTGATGATGAGGTCTCGGAGCCGCGATCTCCATCCCTGTGAGTCCCTCC 1839  
Db 164834 GGGGACACTGATGATGAGGTCTCGGAGCCGCGATCTCCATCCCTGTGAGTCCCTCC 164893  
  
QY 1840 ACCTCAGCTCCCACTCTGCTGCTTTTGGAGCTTAAGCCTGGGATGACCAAAATTCACCCC 1899  
Db 164894 ACCTCAGCTCCCACTCTGCTGCTTTTGGAGCTTAAGCCTGGGATGACCAAAATTCACCCC 164953  
  
QY 1900 AGCTCCTTCATTCACAGGCTGGATGATGAGTGGGATTGAGTCCATGTTATCGGCTCGGTA 1959  
Db 164954 AGCTCCTTCATTCACAGGCTGGATGATGAGTGGGATTGAGTCCATGTTATCGGCTCGGTA 165013  
  
QY 1960 CTCAACACACCAAGTTTCATCCGAGGAAATGTCCCGGAGTGGATGAGCTCACAATGC 2019  
Db 165014 CTCAACACACCAAGTTTCATCCGAGGAAATGTCCCGGAGTGGATGAGCTCACAATGC 165073  
  
QY 2020 TGAGGAACACCCAGCTCTGGACAGAGTTCTTATAAATGATAAATAGGCTCAGAAACCA 2079  
Db 165074 TGAGGAACACCCAGCTCTGGACAGAGTTCTTATAAATGATAAATAGGCTCAGAAACCA 165133  
  
QY 2080 CTGCATTTGACCTGCTGTACAGACTGCCACACACTGCTGACCTGCCCTAGCGAGCAGACA 2139  
Db 165134 CTGCATTTGACCTGCTGTACAGACTGCCACACACTGCTGACCTGCCCTAGCGAGCAGACA 165193  
  
QY 2140 TCCCTTCTGAGCCATCTGCTGCTCTCTCATTTTCATACCCCAACTGTCCCTTTGTTGA 2199
```

```
Db 165194 TCCCTTCTGAGCCATCTGCTGCTCTCTCATTTTCATCCACCAACTGTCCCTTGTTTTGA 165253
QY 2200 TCAATGGGACCGCCACTGCCAGGAGCAGCTTTAGGGCTCTCAGTTCAAACTGAAGGA 2259
Db 165254 TCAATGGGACCGCCACTGCCAGGAGCAGCTTTAGGGCTCTCAGTTCAAACTGAAGGA 165313
QY 2260 CAGTTGAACCTCAGATGGGGTTTCATGTGGGATTTCTGGGAGCTTCTGGGAAATTCAGTTGGA 2319
Db 165314 CAGTTGAACCTCAGATGGGGTTTCATGTGGGATTTCTGGGAGCTTCTGGGAAATTCAGTTGGA 165373
QY 2320 GTCAAGTCAGATGCTCTCAAGGACCCCTCGGGCTCAGAGCCCTAAAGTGGGCCCTGGTG 2379
Db 165374 GTCAAGTCAGATGCTCTCAAGGACCCCTCGGGCTCAGAGCCCTAAAGTGGGCCCTGGTG 165433
QY 2380 AAGCAGGGTGTCTCGCTGCTCCACTTCCCAAGCCTGAGCAGCTCATCTTCATTTGAATGT 2439
Db 165434 AAGCAGGGTGTCTCGCTGCTCCACTTCCCAAGCCTGAGCAGCTCATCTTCATTTGAATGT 165493
QY 2440 CTCAATTTGGCGGAGGAACAACTGAACCTTTGCTGTTTGTCTTTAGCCCTTCAGTTGCTCC 2499
Db 165494 CTCAATTTGGCGGAGGAACAACTGAACCTTTGCTGTTTGTCTTTAGCCCTTCAGTTGCTCC 165553
QY 2500 GCTGCCCTCCACCAGAGGTTTGGCGAGCCTGTGTTGCAGGGTTGTATAAAACCAAGGT 2559
Db 165554 GCTGCCCTCCACCAGAGGTTTGGCGAGCCTGTGTTGCAGGGTTGTATAAAACCAAGGT 165613
QY 2560 ACTTCGTTAGTTTGGCCATTACGCCATGTCAGCTGACATGCAAGTAATCTTTCCTCCT 2619
Db 165614 ACTTCGTTAGTTTGGCCATTACGCCATGTCAGCTGACATGCAAGTAATCTTTCCTCCT 165673
QY 2620 AATTATAGAATGATTTTCTTTTAAATTTTACTTTTACCAGACTTTACTTTTGACTCAG 2679
Db 165674 AATTATAGAATGATTTTCTTTTAAATTTTACTTTTACCAGACTTTACTTTTGACTCAG 165733
QY 2680 AGAAGAGCCCTCAGATGGCTGTGTACATATAATGTTGACTAACTCTTAAATAAAAAA 2739
Db 165734 AGAAGAGCCCTCAGATGGCTGTGTACATATAATGTTGACTAACTCTTACTTTGAAC 165793
QY 2740 AAAAAAA 2747
Db 165794 CAGGAAA 165801

RESULT 5
E38420
LOCUS
DEFINITION Novel polypeptide.
ACCESSION E38420
VERSION E38420.1 GI:18626994
KEYWORDS JP 2000245464-A/2.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 10562)
Narimatsu, H., Isshiki, S., Togayauchi, A. and Sasaki, K.
Novel polypeptide
Patent: JP 2000245464-A 2 12-SEP-2000;
KYOWA HAKKO KOGYO CO LTD
OS Homo sapiens (human)
PN JP 2000245464-A/2
PD 12-SEP-2000
PF 25-FEB-1999 JP 1999047571
PR
PI HISASHI NARIMATSU, SOICHIRO ISSHIKI, AKIRA TOGAYAUCHI, PI
KATSUTOSHI SASAKI
PC C12N15/09, A01K67/027, C12N1/21, C12N5/10, C12N9/10, C12P19/00, PC
C12P21/02,
PC C12P21/08, C12Q1/68, G01N33/53///(C12N1/21, C12R1/185), (C12N5/10,
C12R1.91),
PC (C12P21/02, C12R1.185), (C12P21/02, C12R1.91), C12N5/00, C12N5/00,
C12N5/00, C12R1.91)
```

```
CC Key Location/Qualifiers
FH Promoter (1) (5000)
FT exon (5001) (5140)
FT exon (5001) (5273)
FT exon (5459) (5567)
FT exon (7427) (7586)
FT exon (8234) (10562).
FEATURES
source
1..10562
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2610 a 2415 c 2574 g 2963 t
ORIGIN
Query Match 83.9%; Score 2329.4; DB 6; Length 10562;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2330; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 400 AAATGGCTTTCCCGAAGATGAGATTGATGATATATTTGCTTCTGCTTCTGGGGGCTCTTTT 459
Db 8232 AGATGGCTTTCCCGAAGATGAGATTGATGATATATTTGCTTCTGCTTCTGGGGGCTCTTT 8291
QY 460 GTTTGTATTTTAGCATGTACAGTCTAAATCCCTTCAAGAAACAGTCTCTTTGTTTACAAGA 519
Db 8292 GTTTGTATTTTAGCATGTACAGTCTAAATCCCTTCAAGAAACAGTCTCTTTGTTTACAAGA 8351
QY 520 AAGACGGGAACCTTCTTAAGCTCCAGATACAGATGAGGAGGAGACACCTTCCCTTCCCTCG 579
Db 8352 AAGACGGGAACCTTCTTAAGCTCCAGATACAGATGAGGAGGAGACACCTTCCCTTCCCTCG 8411
QY 580 TCCTGCTGCTGACCTCATCCCAACAAACAGTTGGCTGAGCGCATGCCATCCGGCAGACGT 639
Db 8412 TCCTGCTGCTGACCTCATCCCAACAAACAGTTGGCTGAGCGCATGCCATCCGGCAGACGT 8471
QY 640 GGGGAAAGAGAGATGCTGAAGGAAAGACAGCTGAAGACATCTTCTCTCTGGGGACCA 699
Db 8472 GGGGAAAGAGAGATGCTGAAGGAAAGACAGCTGAAGACATCTTCTCTCTGGGGACCA 8531
QY 700 CCAGCAGTGCAGCGGAACGAAGAGGTGGACCGAGGAGCGACGACGAGGAGGACATTA 759
Db 8532 CCAGCAGTGCAGCGGAACGAAGAGGTGGACCGAGGAGCGACGACGAGGAGGACATTA 8591
QY 760 TCCAGAGGATTTCTTAGACGTCTATTACAATCTGACCTTGAGACCATGATGGGCATAG 819
Db 8592 TCCAGAGGATTTCTTAGACGTCTATTACAATCTGACCTTGAGACCATGATGGGCATAG 8651
QY 820 AATGGGTCATCGCTTTTGTCTCAGCGCGGCTTTGTGATGAAACAGACTCAGACATGT 879
Db 8652 AATGGGTCATCGCTTTTGTCTCAGCGCGGCTTTGTGATGAAACAGACTCAGACATGT 8711
QY 880 TCATCAATGTTGACTATCTGACTGAATGCTTCTGAAGAAACAGAAACACAGGTTT 939
Db 8712 TCATCAATGTTGACTATCTGACTGAATGCTTCTGAAGAAACAGAAACACAGGTTT 8771
QY 940 TCATGCTCTCTTCAAACTCAATGAGTTTCCCATCAGGACGACCATTCAGCAGTGGTTG 999
Db 8772 TCATGCTCTCTTCAAACTCAATGAGTTTCCCATCAGGACGACCATTCAGCAGTGGTTG 8831
QY 1000 TCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGGCACCCTGCTAG 1059
Db 8832 TCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGGCACCCTGCTAG 8891
QY 1060 TGTTTCTGGGACGTGGCGAGTCAATGTGTACAATGTCTCAAGAGCGCTCCCATACATTA 1119
Db 8892 TGTTTCTGGGACGTGGCGAGTCAATGTGTACAATGTCTCAAGAGCGCTCCCATACATTA 8951
QY 1120 AACTGGAAGACGTGTTTGTGGGCTCTGCCCTCGAAAGGCTCAACATCAGATGGAGGAGC 1179
Db 8952 AACTGGAAGACGTGTTTGTGGGCTCTGCCCTCGAAAGGCTCAACATCAGATGGAGGAGC 9011
QY 1180 TCCACTCCCGACCGACCTTTTTCAGGGGCTTTACGGCTTCTCCGTATGCTCTTTCAGGA 1239
|||||
```

Db 9012 TCACACTCCAGCGGACGCTTTTTCAGGGGGCTTACGCTTCTCCGTATGCTCTTCAGGA 9071
QY 1240 GGATCGTGGCTGCCACTTCATCAAGCCTCGGACTCTCTTGGACTACTGGCAGGCTCTAG 1299
Db 9072 GGATCGTGGCTGCCACTTCATCAAGCCTCGGACTCTCTTGGACTACTGGCAGGCTCTAG 9131
QY 1300 AGAATCCCGGGGGAAGATTGCGGCTGTCTAGGGAGCCAGAGGCACATCCGGAC 1359
Db 9132 AGAATCCCGGGGGAAGATTGCGGCTGTCTAGGGAGCCAGAGGCACATCCGGAC 9191
QY 1360 AGATTTCAGATAACCCGCTGGGATAGTTTTCGTAGATTGGAAGAGGGGGGACAG 1419
Db 9192 AGATTTCAGATAACCCGCTGGGATAGTTTTCGTAGATTGGAAGAGGGGGGACAG 9251
QY 1420 AGGATGCTGTCTTCAGTGTGAATTCACGCCAGCAAGTGTCTGATGAAGTCACTG 1479
Db 9252 AGGATGCTGTCTTCAGTGTGAATTCACGCCAGCAAGTGTCTGATGAAGTCACTG 9311
QY 1480 ATTAGTTCCCACTGTGTGCCCCAGGCAATAATAGGCCGCTCTTGGGCACGCACACTCT 1539
Db 9312 ATTAGTTCCCACTGTGTGCCCCAGGCAATAATAGGCCGCTCTTGGGCACGCACACTCT 9371
QY 1540 TCATACTAAGTGTTCGACATACACCTGGATTTCGATTTCAGGGGTTCAGTATCCTATGA 1599
Db 9372 TCATACTAAGTGTTCGACATACACCTGGATTTCGATTTCAGGGGTTCAGTATCCTATGA 9431
QY 1600 CATGATGGGTTCACCATCTAATTTTACAGCAAGACACAGCTGCGAGAGGTACA 1659
Db 9432 CATGATGGGTTCACCATCTAATTTTACAGCAAGACACAGCTGCGAGAGGTACA 9491
QY 1660 GAAACTTGTCCCAAGGCTCACAGCCAGTAGGCATAGAGCGGGAATCAAAATCGACACT 1719
Db 9492 GAAACTTGTCCCAAGGCTCACAGCCAGTAGGCATAGAGCGGGAATCAAAATCGACACT 9551
QY 1720 GTCAGAACTGTGGGAGCCCTGACTTGAACCACTCCACAGTGTGCTCCCTTAGGA 1779
Db 9552 GTCAGAACTGTGGGAGCCCTGACTTGAACCACTCCACAGTGTGCTCCCTTAGGA 9611
QY 1780 GGGGACACTGATGATGAGTCTCGGAGCGGCATCTTCATCCCTGTCGAGTCCCTCC 1839
Db 9612 GGGGACACTGATGATGAGTCTCGGAGCGGCATCTTCATCCCTGTCGAGTCCCTCC 9671
QY 1840 ACCTCAGCTCCCACTGCTTGTGCTTTTGGAGCTAAGCCCTGGATGACCAAAATTCACCCC 1899
Db 9672 ACCTCAGCTCCCACTGCTTGTGCTTTTGGAGCTAAGCCCTGGATGACCAAAATTCACCCC 9731
QY 1900 AGCTCTCTTCATTCACAGGCTGATGATGAGTGGGATTTAGTCACTGTTATCGGCTCGGTA 1959
Db 9732 AGCTCTCTTCATTCACAGGCTGATGATGAGTGGGATTTAGTCACTGTTATCGGCTCGGTA 9791
QY 1960 CTCACACAAACCAAGCTTTCATCCGAGGAAATGTCCTCCCGCAGTGGATGACATGC 2019
Db 9792 CTCACACAAACCAAGCTTTCATCCGAGGAAATGTCCTCCCGCAGTGGATGACATGC 9851
QY 2020 TGAGGAACACCCAGCTCTGACAGAGTCTTATAAATGTATAAATAGGCTCAGAAACCA 2079
Db 9852 TGAGGAACACCCAGCTCTGACAGAGTCTTATAAATGTATAAATAGGCTCAGAAACCA 9911
QY 2080 CTGCATTCGACTGCTGTACAGACTGCCACACTGCTGACCTGCCCTAGGAGCAGACA 2139
Db 9912 CTGCATTCGACTGCTGTACAGACTGCCACACTGCTGACCTGCCCTAGGAGCAGACA 9971
QY 2140 TCCCTTCTCAGGCATCTGCTGCTCTCTCATTTTCATCACCACCACTGCTCTTTTGA 2199
Db 9972 TCCCTTCTCAGGCATCTGCTGCTCTCTCATTTTCATCACCACCACTGCTCTTTTGA 10031
QY 2200 TCAATGGGACACGCACTGCCCGAGGAGCACTTTAGGGCTCTCAGTTCAAACTGAAGGA 2259
Db 10032 TCAATGGGACACGCACTGCCCGAGGAGCACTTTAGGGCTCTCAGTTCAAACTGAAGGA 10091
QY 2260 CAGTTGAACCTAGATGGGGTTCATGTGGGATTTCTGGGAGCTTTCTGGGAAATTCAGTTGA 2319
Db 10092 CAGTTGAACCTAGATGGGGTTCATGTGGGATTTCTGGGAGCTTTCTGGGAAATTCAGTTGA 10151

QY 2320 GTCAAGTCAGAGTCTCTCAAGGACCCCTCGGGCTCAGAGCCCTAAAAGTGGCCCTGGTG 2379
Db 10152 GTCAAGTCAGAGTCTCTCAAGGACCCCTCGGGCTCAGAGCCCTAAAAGTGGCCCTGGTG 10211
QY 2380 AAGCAGGCTGGTCTCGGTCCTCACTTCCCAAGCCTGAGCCCAAGCTCATCTTCATTTGAATGT 2439
Db 10212 AAGCAGGCTGGTCTCGGTCCTCACTTCCCAAGCCTGAGCCCAAGCTCATCTTCATTTGAATGT 10271
QY 2440 CTCAATTTGGCCGAGGAACAACACTTGTGTTGCTGTTTGTAGCCCTCAGTTTGTCTCC 2499
Db 10272 CTCAATTTGGCCGAGGAACAACACTTGTGTTGCTGTTTGTAGCCCTCAGTTTGTCTCC 10331
QY 2500 GCTGCTCTTACCCAGAGGTTTGTGCGAGCCTGTGTTGCGAGGTTGTATAAAACCAAGGT 2559
Db 10332 GCTGCTCTTACCCAGAGGTTTGTGCGAGCCTGTGTTGCGAGGTTGTATAAAACCAAGGT 10391
QY 2560 ACTTCGTTAGTTTGGCCATTCAGCCATGTCAGCTGACATGCAAAAGTAATCTTGTCTCT 2619
Db 10392 ACTTCGTTAGTTTGGCCATTCAGCCATGTCAGCTGACATGCAAAAGTAATCTTGTCTCT 10451
QY 2620 AATTATAGAAATGATTTTCTTTTAAATTTTACTTTTACCAGACTTTTACTTTGTACTCAG 2679
Db 10452 AATTATAGAAATGATTTTCTTTTAAATTTTACTTTTACCAGACTTTTACTTTGTACTCAG 10511
QY 2680 AGAAGGCTCATGCTGCTGCATATATAAATGTTGGACTAAACTCTT 2730
Db 10512 AGAAGGCTCATGCTGCTGCATATATAAATGTTGGACTAAACTCTT 10562

RESULT 6
AF145784
LOCUS
DESCRIPTION
Homo sapiens beta1,3 galactosyltransferase-V (B3GALT5) gene,
complete cds.
ACCESSION AF145784
VERSION AF145784.1 GI:6409192
KEYWORDS
SOURCE
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 933)
AUTHORS Zhou,D., Berger,E.G. and Hennet,T.
TITLE Molecular cloning of a human UDP-galactose:GlcNAc6phosphate-3-galactose-4-epimerase gene encoding an O-linked core3-elongation enzyme
JOURNAL Eur. J. Biochem. 263 (2), 571-576 (1999)
MEDLINE 99337698
PUBMED 10406968
REFERENCE 2 (bases 1 to 933)
AUTHORS Zhou,D. and Hennet,T.
TITLE Direct Submission
JOURNAL Submitted (26-APR-1999) Physiology, University of Zurich, Winterthurerstrasse 190, Zurich 8057, Switzerland
FEATURES
source
1. .933
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22"
1. .933
/gene="B3GALT5"
1. .933
/gene="B3GALT5"
/note="glycosyltransferase"
/codon_start=1
/product="beta1,3 galactosyltransferase-V"
/protein_id="AAF07880.1"
/db_xref="GI:6409193"
/translation="MAFPKRLMYICLLVLGALCLVFSYSLNPKFQSVYVKDGNF
LKLPTDCROTTPFLVLLVTSKHLAERMAIRQTGKRWYKGLQKLTFFLLGTTSS
AAETKEVDQESQRHGDIIQKDFLDVYINLTATMMGIEMVHRFCQAAPVMTDSDF

INVLYTELLKKNNRTTRFTFLKLNFFPIROPFSKWFVSKSEYPWDYRPPPCSGTGY
YVFGSDVASQVYNVSKSVPIKLEDFVGLCLERLNLRLLELHSHQPTFFPGGLRFSVC
LFRRIVACHPIKPTLLDYWQALENSRGEDCPV*

BASE COUNT 227 a 232 c 241 g 233 t
ORIGIN

Query Match 33.68; Score 933; DB 9; Length 933;
Best Local Similarity 100.0%; Pred. No. 2.2e-236;
Matches 933; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 402 ATGCTTTCCCGAAGATGAGATTGATATATTTGCCCTCTGGTCTGGGGGCTCTTTGT 461
|||||
Db 1 ATGCTTTCCCGAAGATGAGATTGATATATTTGCCCTCTGGTCTGGGGGCTCTTTGT 60
|||||

QY 462 TTGTATTTTACATGTACAGTCTAAATCTTTCAAAGAACAGTCTTTGTTTACAAGAAA 521
|||||
Db 61 TTGTATTTTACATGTACAGTCTAAATCTTTCAAAGAACAGTCTTTGTTTACAAGAAA 120
|||||

QY 522 GACGGGAACCTTCCCTTAAGCTCCAGATACAGCTGCAGGAGACACCTCCCTTCCTGTC 581
|||||
Db 121 GACGGGAACCTTCCCTTAAGCTCCAGATACAGCTGCAGGAGACACCTCCCTTCCTGTC 180
|||||

QY 582 CTGCTGTGACCTCATCCCAAAACAGTTGGCTGAGCGCATGGCCATCCGGCAGACGTGG 641
|||||
Db 181 CTGCTGTGACCTCATCCCAAAACAGTTGGCTGAGCGCATGGCCATCCGGCAGACGTGG 240
|||||

QY 642 GGGAAAGAGAGGATGGTGAAGGGAAGCAGCTGAAGACATTTCTCTCTCTGGGACCACC 701
|||||
Db 241 GGGAAAGAGAGGATGGTGAAGGGAAGCAGCTGAAGACATTTCTCTCTCTGGGACCACC 300
|||||

QY 702 AGCAGTCCAGCGGAACGAAGAGCGTGACAGAGAGCGACGACGAGGGGACATATC 761
|||||
Db 301 AGCAGTCCAGCGGAACGAAGAGCGTGACAGAGAGCGACGACGAGGGGACATATC 360
|||||

QY 762 CAGAAGGATTTCTTACAGCTCTATTACAATCTGACCTGAAGACCATGATGGGCATAGAA 821
|||||
Db 361 CAGAAGGATTTCTTACAGCTCTATTACAATCTGACCTGAAGACCATGATGGGCATAGAA 420
|||||

QY 822 TGGTCCATCGCTTTTCTCAGCGCGGCTTTGTGATGAAACAGACTCAGACATGTTTC 881
|||||
Db 421 TGGTCCATCGCTTTTCTCAGCGCGGCTTTGTGATGAAACAGACTCAGACATGTTTC 480
|||||

QY 882 ATCAATGTTGACTATCTGACTGAATGCTTCTGAAGAAAACAGAACACAGGTTTTC 941
|||||
Db 481 ATCAATGTTGACTATCTGACTGAATGCTTCTGAAGAAAACAGAACACAGGTTTTC 540
|||||

QY 942 ACTGGCTTCTTGAACATCAATGAGTTTCCCATCAGGCGACCATTCAGCAAGTGTGTTGTC 1001
|||||
Db 541 ACTGGCTTCTTGAACATCAATGAGTTTCCCATCAGGCGACCATTCAGCAAGTGTGTTGTC 600
|||||

QY 1002 AGTAATCTGAATATCCGTGGGACAGTACCCACCATTCGTCTCCGGCACCGGTACGTG 1061
|||||
Db 601 AGTAATCTGAATATCCGTGGGACAGTACCCACCATTCGTCTCCGGCACCGGTACGTG 660
|||||

QY 1062 TTTTCTGGCAGCTGGGAGTCAAGTCTCAAAATGCTCTCCAGAGCGTCCCATACATATAA 1121
|||||
Db 661 TTTTCTGGCAGCTGGGAGTCAAGTCTCAAAATGCTCTCCAGAGCGTCCCATACATATAA 720
|||||

QY 1122 CTGGAAGACGTGTTGTGGGGCTCTGCTCTGAAAGGCTGAACATCAGATTGGAGGAGCTC 1181
|||||
Db 721 CTGGAAGACGTGTTGTGGGGCTCTGCTCTGAAAGGCTGAACATCAGATTGGAGGAGCTC 780
|||||

QY 1182 CACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCTCCGATGCCCTTTCAGAGG 1241
|||||
Db 781 CACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCTCCGATGCCCTTTCAGAGG 840
|||||

QY 1242 ATCGTGGCTGCCACCTTCATCAGCCCTCGACCTCTCTTGGACTACTGGCAGGCTCTAGAG 1301
|||||
Db 841 ATCGTGGCTGCCACCTTCATCAGCCCTCGGACTCTCTTGGACTACTGGCAGGCTCTAGAG 900
|||||

QY 1302 AATTCCCGGGGGAAGATTGTCCGCCTGTCTGA 1334
|||||
Db 901 AATTCCCGGGGGAAGATTGTCCGCCTGTCTGA 933
|||||

RESULT 7
HSA6078 Homo sapiens beta3gal-T6 gene. 933 bp DNA linear PRI 11-MAY-2000
LOCUS
DEFINITION
ACCESSION AJ006078
VERSION
KEYWORDS beta-1,3-galactosyltransferase; beta3gal-T6 gene.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Amado, M., Carneiro, F. and Clausen, H.
TITLE Cloning and expression of two beta-1,3-galactosyltransferases:
beta3gal-T5 and beta3gal-T6
JOURNAL Unpublished
AUTHORS Amado, M.
TITLE Direct Submission
JOURNAL Submitted (11-MAY-1998) Amado M., Department of Oral Diagnostics,
Royal dental School, Norre Allie 20, 2200 Copenhagen, DENMARK
FEATURES
source
1. .933 Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene
1. .933
/gene="beta3gal-T6"
CDS
1. .933
/gene="beta3gal-T6"
/product="beta-1,3-galactosyltransferase"
/protein_id="CAB01547.1"
/db_xref="GI:779923"
/translation="MAFPKRLMYICLLVGLALCLYYSMNSLNPFEQSFYVKDGNF
LKLPDTCROTTPPLVLLVTSSSHQLAERAIROTWGERVTKVQKQLKTFLLGTTSS
AAETKEVDQESORHCDIIQKDFLDVYNLTWKMGIEWHVFQPOAAAFVHKVTDSDMF
INVLYTELLKKNNRTTRFTFLKLNFFPIROPFSKWFVSKSEYPWDYRPPPCSGTGY
YVFGSDVASQVYNVSKSVPIKLEDFVGLCLERLNLRLLELHSHQPTFFPGGLRFSVC
LFRRIVACHPIKPTLLDYWQALENSRGEDCPV*"
BASE COUNT 229 a 234 c 240 g 229 t 1 others
ORIGIN

Query Match 33.38; Score 924; DB 9; Length 933;
Best Local Similarity 99.48; Pred. No. 5.4e-234;
Matches 927; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 402 ATGCTTTCCCGAAGATGAGATTGATATATTTGCCCTCTGGTCTGGGGGCTCTTTGT 461
|||||
Db 1 ATGCTTTCCCGAAGATGAGATTGATATATTTGCCCTCTGGTCTGGGGGCTCTTTGT 60
|||||

QY 462 TTGTATTTTACATGTACAGTCTAAATCTTTCAAAGAACAGTCTTTGTTTACAAGAAA 521
|||||
Db 61 TTGTATTTTACATGTACAGTCTAAATCTTTCAAAGAACAGTCTTTGTTTACAAGAAA 120
|||||

QY 522 GACGGGAACCTTCCCTTAAGCTCCAGATACAGCTGCAGGAGACACCTCCCTTCCTGTC 581
|||||
Db 121 GACGGGAACCTTCCCTTAAGCTCCAGATACAGCTGCAGGAGACACCTCCCTTCCTGTC 180
|||||

QY 582 CTGCTGTGACCTCATCCCAAAACAGTTGGCTGAGCGCATGGCCATCCGGCAGACGTGG 641
|||||
Db 181 CTGCTGTGACCTCATCCCAAAACAGTTGGCTGAGCGCATGGCCATCCGGCAGACGTGG 240
|||||

QY 642 GGGAAAGAGAGGATGGTGAAGGGAAGCAGCTGAAGACATTTCTCTCTCTGGGACCACC 701
|||||
Db 241 GGGAAAGAGAGGATGGTGAAGGGAAGCAGCTGAAGACATTTCTCTCTCTGGGACCACC 300
|||||

QY 702 AGCAGTCCAGCGGAACGAAGAGCGTGACAGAGAGCGACGACGAGGGGACATATC 761
|||||
Db 301 AGCAGTCCAGCGGAACGAAGAGCGTGACAGAGAGCGACGACGAGGGGACATATC 360
|||||

QY 762 CAGAAGGATTTCTTACAGCTCTATTACAATCTGACCTTGAAGACCATGATGGGCATAGAA 821
|||||


```
QY 1120 AACTGGAAGAGCTGTTTCTGGGGCTCTCCCTCGAAAGGCTGAACATCAGATTGAGGAGC 1179
|||||
Db 1393 AACTGGAAGAGCTGTTTCTGGGGCTCTCCCTCGAAAGGCTGAACATCAGATTGAGGAGC 1452

QY 1180 TCACACTCCAGCGACACCTTTTTCACAGGGGCTTACGCTTCTCCGTATGCGCTCTTCAGGA 1239
|||||
Db 1453 TCACACTCCAGCGACACCTTTTTCACAGGGGCTTACGCTTCTCCGTATGCGCTCTTCAGGA 1512

QY 1240 GGATCGTGGGCTGCCACTTCATCAAGCCTCGGACTCTCTTGGACTACTGGCAGGCTCTAG 1299
|||||
Db 1513 GGATCGTGGGCTGCCACTTCATCAAGCCTCGGACTCTCTTGGACTACTGGCAGGCTCTAG 1572

QY 1300 AGAA 1303
|||||
Db 1573 AGAA 1576

RESULT 9
AB041415 1570 bp DNA linear PRI 13-APR-2000
LOCUS
DEFINITION
Pan paniscus beta1,3-GalT 5 gene for UDP-Gal:GlcNAc
beta1,3-galactosyltransferase 5, partial cds.
ACCESSION
AB041415
VERSION
AB041415.1 GI:7593024
KEYWORDS
SOURCE
Pan paniscus (isolate: bonobo-05) DNA.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
REFERENCE
1 (bases 1 to 1570)
AUTHORS
Liu, Y. and Saitou, N.
JOURNAL
Silver Project
PUBLISHED ONLY in DataBase (2000)
REFERENCE
2 (bases 1 to 1570)
AUTHORS
Liu, Y. and Saitou, N.
JOURNAL
Direct Submission
Submitted (11-APR-2000) Naruya Saitou, National Institute of
Genetics, Laboratory of Evolutionary Genetics; 1111 Yata, Mishima,
Shizuoka 411-8540, Japan (E-mail: nsaitou@genetics.nig.ac.jp,
URL: http://sayer.lab.nig.ac.jp/~silver/, Tel: 81-559-81-6790,
Fax: 81-559-81-6789)
FEATURES
Location/Qualifiers
1..1570
/organism="Pan paniscus"
/isolate="bonobo-05"
/db_xref="taxon:9597"
/note="human sequence used for primer design based on Acc#
AB020337"
<1..20
/exon
/number=3
21..667
/intron
/number=3
668..1570
/gene="beta1,3-GalT 5"
668..>1570
/gene="beta1,3-GalT 5"
/codon_start=1
/product="UDP-Gal:GlcNAc beta1,3-galactosyltransferase 5"
/db_xref="GI:7593025"
/translacion="MAFPKRLMYVCLLVGLALCLYFSMYSLNLFKEQSFVYKKDGNF
LKLPDTRQTPFPFLVLIVTSSHQLAERMAIRQWTKERTVKGQKLTFFLLGTTSS
AAETKEVDQESQRHGDIIQKDFLDGYNLTLKTMGMIEWHRFCPOAAVFMKTDSDMF
INVDVLTLLKKNRTTRFTFKLNLNFFIRQPFKSVKSEYVWDYRPPPCSGTG
VVFSDVASQVYNVSEVPYIKLEDVVFVGLCLERLNLRLLELHQSQPTFPFGLRFSVC
REFRIVACHFIKPRTLDDYWALEN"
668..>1570
/exon
/gene="beta1,3-GalT 5"
/number=4
408 a 372 c 392 g 398 t
BASE COUNT
ORIGIN
```

```
Query Match 31.8%; Score 882.6; DB 9; Length 1570;
Best Local Similarity 98.5%; Pred. No. 5.8e-223;
Matches 891; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 400 AAATGGGCTTTCCCGAAGATGATGATGATATATTGCTTCTGCTGCTGGGGGCTCTTT 459
|||||
Db 666 AGATGGGCTTTCCCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 725

QY 460 GTTTGTATTTAGCATGTACAGTCTAAATCTTTTCAAGAAGACAGTCTCTTTGTTTACAAGA 519
|||||
Db 726 GTTTGTATTTAGCATGTACAGTCTAAATCTTTTCAAGAAGACAGTCTCTTTGTTTACAAGA 785

QY 520 AAGACGGGAACCTTCTTAAAGCTCCAGATACAGACTGCGAGGAGACACCTCCCTTCCTCG 579
|||||
Db 786 AAGATGGGAACCTTCTTAAAGCTCCAGATACAGACTGCGAGGAGACACCTCCCTTCCTCG 845

QY 580 TCCTGCTGGTGACCTCATCCCAACAAAGTGTGGCTGAGCGCATGCCATCCGGCAGAGCT 639
|||||
Db 846 TCCTGCTGGTGACCTCATCCCAACAAAGTGTGGCTGAGCGCATGCCATCCGGCAGAGCT 905

QY 640 GGGGAAAGAGAGAGTGGTGAAGGAAAGAGAGCTGAAGACATTTCTTCTCTCTGGGAGCA 699
|||||
Db 906 GGGGAAAGAGAGAGAGTGGTGAAGGAAAGAGAGCTGAAGACATTTCTTCTCTCTGGGAGCA 965

QY 700 CCAGCAGTGCAGCGGAACAAAGAGGTGGACAGGAGCGACGACACCGGAGCATTA 759
|||||
Db 966 CCAGCAGTGCAGCGGAACAAAGAGGTGGACAGGAGCGACGACCGGAGCATTA 1025

QY 760 TCCAGAAGGATTTCTTAGACGCTCTATTACAATCTGACCCCTGAAGACCATGATGGCATAG 819
|||||
Db 1026 TCCAGAAGGATTTCTTAGACGCTCTATTACAATCTGACCCCTGAAGACCATGATGGCATAG 1085

QY 820 AATGGGTCATCGCTTTTGTCTCAGCGGGGCTTTGTGATGATGATGATGATGATGATGATGAT 879
|||||
Db 1086 AATGGGTCATCGCTTTTGTCTCAGCGGGGCTTTGTGATGATGATGATGATGATGATGATGAT 1145

QY 880 TCATCAATGTTGACTATCTGACTGAACCTGCTTCTGAAGAAAACAGAACACACAGGTTTT 939
|||||
Db 1146 TCATCAATGTTGACTATCTGACTGAACCTGCTTCTGAAGAAAACAGAACACACAGGTTTT 1205

QY 940 TCACTGGCTTCTTGAAGTCAATGAGTTTCCCATCAGGACGACCATTCAGCAAGTGGTTTG 999
|||||
Db 1206 TCACTGGCTTCTTGAAGTCAATGAGTTTCCCATCAGGACGACCATTCAGCAAGTGGTTTG 1265

QY 1000 TCAGTAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGGCACCCTGCTAG 1059
|||||
Db 1266 TCAGTAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGGCACCCTGCTAG 1325

QY 1060 TGTTTTCTGGCGAGTGGCGAGTCAAGTGTACAATGTCTCCAAGAGCGCTCCCATACATTA 1119
|||||
Db 1326 TGTTTTCTGGCGAGTGGCGAGTCAAGTGTACAATGTCTCCAAGAGCGCTCCCATACATTA 1385

QY 1120 AACTGGAAGAGGCTTTTGTGGGCTCTGCTTCGAAAGGCTGAACATCAGATTGAGGAGC 1179
|||||
Db 1386 AACTGGAAGAGGCTTTTGTGGGCTCTGCTTCGAAAGGCTGAACATCAGATTGAGGAGC 1445

QY 1180 TCCACTCCAGCGACCTTTTCCAGGGGCTTACGCTTCTCCGTATGCTCTCTTCAGGA 1239
|||||
Db 1446 TCCACTCCAGCGACCTTTTCCAGGGGCTTACGCTTCTCCGTATGCTCTCTTCAGGA 1505

QY 1240 GGATCGTGGGCTGCCACTTCATCAAGCCTCGGACTCTCTTGGACTACTGGCAGGCTCTAG 1299
|||||
Db 1506 GGATCGTGGGCTGCCACTTCATCAAGCCTCGGACTCTCTTGGACTACTGGCAGGCTCTAG 1565

QY 1300 AGAA 1304
|||||
Db 1566 AGAA 1570
```

```
RESULT 10
AB041413
LOCUS
DEFINITION Homo sapiens beta1,3-GalT 5 gene for UDP-Gal:GlcNAc
```


* 146379 146478: gap of unknown length
* 146479 161686: contig of 15208 bp in length
* 161687 161786: gap of unknown length
* 161787 176149: contig of 14363 bp in length
* 176150 176249: gap of unknown length
* 176250 196900: contig of 20651 bp in length.

FEATURES

Location/Qualifiers
source

1. 196900
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP21-467L12"

BASE COUNT 51456 a 43823 c 44608 g 53666 t 3347 others
ORIGIN

Query Match 19.5%; Score 542.2; DB 2; Length 196900;

Best Local Similarity 73.8%; Pred. No. 2.9e-132;

Matches 734; Conservative 0; Mismatches 248; Indels 13; Gaps 3;

```
QY 400 AATGCTTTCCGAGATGAGATTCATGATATATTTGCCCTTCTGGTTCGGGGCTCTTT 459
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102456 AGATGCTTCACATGAAGAAGGCTAGTTTACGCCCTCCATTCTCATGATGGCGCACTCT 102515
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 460 GTTGTATTATTAGCATGTACAGTCTAAATCCCTTTTCAAGAAGCATGCTCTTTTACAAGA 519
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102516 GCTTGACTTTCAGCATGGA-----TCTTTTCAGAACTCCCGTTGTTTTTAAAG 102566
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 520 AAG---ACGGAACTTCTTAAGTCCCAAGATACAGACTGCAGGCAGACACCTCCCTTCC 576
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102567 AAGTCACGGGAAGTTCCTTCAGATTCGGATATAGACTGCAAGCAGAGCGCCCTTTC 102626
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 577 TCGTCTGCTGGTACCTATCCCAAAACAGTTGGCTGAGCGCATGGCCATCCGCGCAGA 636
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102627 TGTGCTGCTGGTGGTCACTTCACAGCAGCTGGCCGCTCGCATGGCCATCCGCAAGA 102686
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 637 CGTGGGGGAAAGAGAGATGTGAAGGAAAGCAGCTGAAGACATCTTCTCTCTCGGGGA 696
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102687 CGTGGGTAGAGAGACATCTGTGCAGGGCCACAGGTGAGGACCTTCTCTCTCGGGA 102746
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 697 CCACAGCAGTCCGCGGAAAGAAAGAGGTGGACCAGAGAGCCAGCAGCGGGGACA 756
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102747 CTTCCACAGACCCGAGGAGATGGAGCCACAACTCTGGAGCGGAGCAGCAGCGGACA 102806
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 757 TTATCCAGAAGGATTTCTTAGAGCTCTATTACAATCTGACCTTGAAGACCATGATGSGCA 816
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102807 TTATCCAGAAGGATTTCAAGGATGCTTCTTCACTGACCTTGAAGACCATGATGGGTA 102866
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 817 TAGATGGTCCATCGCTTTTGTCTCAGCGCGGCTTGTGATGAAACAGACATCAGACA 876
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102867 TGAATGGTCTACCACTTTTGTCTCAGACAGCTTACGCTGATGAAACGGACTCTGACA 102926
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 877 TGTTCATCAATCTGACTATCTGACTGAATCTGCTTCTGAAGAAAACAGAACACACAGGT 936
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102927 TGTGTGAATGTGGCTATCTGACGGAACTGCTGCTAAAGAAAACAAACAGCAGGT 102986
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 937 TTTTCACTGGCTTTCTGAAACTCAATGAGTTTCCCATCAGGAGCCATTCAGCAAGTGT 996
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102987 TCTTCACAGGCTACATAAAGCCCGACGACTTTCCCATCCGCGAGAGTTCAACAAGTGT 103046
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 997 TTGTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGCGACCGGT 1056
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 103047 TTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTTGTCTCCGCTACTGGT 103106
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1057 AGCTGTTTCTCGGACGCTGGGAGTCAGGTGTACAATGTTCTCAAGAGCGTCCCATACA 1116
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 103107 ATGTCTTTTCCAGCAGCTGGCGATCCAAAGTATACAACCTCAGAGAGCGTTCCGTTCA 103166
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1117 TFAAATCGAAGCAGCTGTTTGGGGCTGCTGCTCGAAGGCTGACATCATGATGGAGG 1176
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 103167 TCAAGCTGAGGATGTGTTTGTGGGCTGCTGTTGGCCAGGTAAAGATCCCGCGCGAGG 103226
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1177 AGCTCCACTCCCGAGCGGACCTTTTTTCCAGGGGGCTTAGCGTTCTCCGCTATGCCCTTCA 1236
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 103227 AGCTGCACACAAACAGACCTTCTTCCCTGGCGGTTTACGCTTTTCCGCTGGCGCTTC 103286
QY 1237 GGAGGATCGGGCTCCCACTTTCATCAAGCCTCGGACTCTCTTGGACTACTGCGAGGCTC 1296
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 103287 AGAAAATTGGGCATGCCATTTTATGAAGCCCGAGGACCTGCTCACTTATGCGAACAC 103346
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1297 TAGAGAAATTCGGGGGGAAGATTTCGGCCTGTCTGAGGGAGCCAGAGGACATCCG 1356
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 103347 TGCAGAACTCGAAAGAACAGGACTGCTCTGCTGAGGGACCTTGGGGCACCT-AG 103405
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1357 GACAAGTTTCAGATAACCCCGTGGGATAGTTTTTG 1391
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 103406 GACAAAATCTTGATAACCTTTGCTGATGATTGCTG 103440
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

AC120346 175861 bp DNA linear HTG 09-MAY-2002
Mus musculus clone RP23-147E11, WORKING DRAFT SEQUENCE, 22 ordered
pieces.
AC120346 3 GI:20514894
HTG: HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
Mus musculus.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 175861)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Mus musculus, clone RP23-147E11
Unpublished
2 (bases 1 to 175861)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukhgalter,B., Brown,A., Camarato,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Larocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,L.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Willson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zaloun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (06-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 175861)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukhgalter,B., Brown,A., Camarato,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Larocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,L.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
```

Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (09-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 9, 2002 this sequence version replaced gi:20503171.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE JOURNAL

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WtBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence-submissions@genome.wi.mit.edu
----- Project Information
Center project name: L20921
Center clone name: 147_E.11
----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 169879 bases at least Q40
Consensus quality: 172188 bases at least Q30
Consensus quality: 173133 bases at least Q20
Insert size: 176000; agarose-fp
Insert size: 173761; sum-of-contigs
Quality coverage: 8.3 in Q20 bases; agarose-fp
Quality coverage: 8.4 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 22 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
This sequence will be replaced
by the finished sequence as soon as it is available and the accession number will be preserved.

1 28835: contig of 28835 bp in length
28836 28935: gap of 100 bp
28936 29562: contig of 627 bp in length
29563 29662: gap of 100 bp
29663 30328: contig of 666 bp in length
30329 30428: gap of 100 bp
30429 32005: contig of 1577 bp in length
32006 32105: gap of 100 bp
32106 33128: contig of 1023 bp in length
33129 33228: gap of 100 bp
33229 34512: contig of 1284 bp in length
34513 34612: gap of 100 bp
34613 36079: contig of 1467 bp in length
36080 36179: gap of 100 bp
36180 37483: contig of 1304 bp in length
37484 37584: gap of 100 bp
37584 39166: contig of 1583 bp in length
39167 39267: gap of 100 bp
39267 41065: contig of 1799 bp in length
41066 41166: gap of 100 bp
41166 43203: contig of 2038 bp in length
43204 43303: gap of 100 bp
43304 47489: contig of 4186 bp in length
47490 47589: gap of 100 bp
47590 50604: contig of 3015 bp in length
50605 50704: gap of 100 bp
50705 56370: contig of 5666 bp in length
56371 56470: gap of 100 bp
56471 64079: contig of 7609 bp in length
64080 64179: gap of 100 bp

* 64180 72619: contig of 8440 bp in length
* 72620 72719: gap of 100 bp
* 72720 81566: contig of 8847 bp in length
* 81567 81666: gap of 100 bp
* 81667 96774: contig of 15108 bp in length
* 96775 96874: gap of 100 bp
* 96875 119647: contig of 22773 bp in length
* 119648 119747: gap of 100 bp
* 119748 147772: contig of 28025 bp in length
* 147773 147872: gap of 100 bp
* 147873 174557: contig of 26685 bp in length
* 174558 174657: gap of 100 bp
* 174658 175861: contig of 1204 bp in length.

FEATURES

Location/Qualifiers
1..175861
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-147E11"
/clone_lib="RPC1-23 Female Mouse BAC"
1..28835
/note="assembly_fragment
clone_end:SP6
vector_side:left"
28936..29562
/note="assembly_fragment"
29663..30328
/note="assembly_fragment"
30429..32005
/note="assembly_fragment"
32106..33128
/note="assembly_fragment"
33229..34512
/note="assembly_fragment"
34613..36079
/note="assembly_fragment"
36180..37483
/note="assembly_fragment"
37584..39166
/note="assembly_fragment"
39267..41065
/note="assembly_fragment"
41166..43203
/note="assembly_fragment"
43304..47489
/note="assembly_fragment"
47590..50604
/note="assembly_fragment"
50705..56370
/note="assembly_fragment"
56471..64079
/note="assembly_fragment"
64180..72619
/note="assembly_fragment"
72720..81566
/note="assembly_fragment"
81667..96774
/note="assembly_fragment"
96875..119647
/note="assembly_fragment"
119748..147772
/note="assembly_fragment"
147873..174557
/note="assembly_fragment"
174658..175861
/note="assembly_fragment
clone_end:T7
vector_side:right"
BASE COUNT 48903 a 39540 c 39029 g 46286 t 2103 others
ORIGIN

Query Match 19.3%; Score 534.4; DB 2; Length 175861;
Best Local Similarity 73.4%; Pred. No. 3,4e-130;
Matches 728; Conservative 0; Mismatches 251; Indels 13; Gaps 3;

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 14:50:38 ; Search time 461.132 Seconds
(without alignments)
13552.085 Million cell updates/sec

Title: US-09-914-152-2
Perfect score: 2775
Sequence: 1 gtgaattctcttctctgc.....aaaaaaaaaaaaaaaaaaaaa 2775

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*

| | |
|-----|--|
| 1: | /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:* |
| 2: | /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:* |
| 3: | /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:* |
| 4: | /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:* |
| 5: | /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:* |
| 6: | /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:* |
| 7: | /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:* |
| 8: | /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:* |
| 9: | /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:* |
| 10: | /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:* |
| 11: | /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:* |
| 12: | /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:* |
| 13: | /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:* |
| 14: | /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:* |
| 15: | /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:* |
| 16: | /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:* |
| 17: | /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:* |
| 18: | /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:* |
| 19: | /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:* |
| 20: | /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:* |
| 21: | /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:* |
| 22: | /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:* |
| 23: | /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:* |
| 24: | /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:* |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|------------------------------|
| 1 | 2775 | 100.0 | 2775 | 21 | AAA93875 Human beta-1,3 gal |
| 2 | 2329.4 | 83.9 | 10562 | 21 | AAA93876 Human beta3Gal-T5 |
| 3 | 928.6 | 33.5 | 1011 | 21 | AAA27959 Human beta3Gal-T5 |
| 4 | 163.4 | 5.9 | 1739 | 15 | AAQ67067 Beta-1,3-galactosyl |
| 5 | 133.6 | 4.8 | 3518 | 19 | AAV57905 Bovine butyrophili |
| 6 | 130.4 | 4.7 | 1773 | 20 | AAH87193 Human Dendriac cDN |
| 7 | 130.4 | 4.7 | 1897 | 24 | ABK51201 Human cDNA encodin |
| 8 | 130.4 | 4.7 | 2095 | 21 | AZ65022 Membrane-bound pro |
| 9 | 130.4 | 4.7 | 2095 | 22 | AAS46005 Human DNA encoding |

| | | | | | |
|----|-------|-----|--------|----|-----------------------------|
| 10 | 130.4 | 4.7 | 2095 | 22 | AAF92075 Human PRO1074 cDNA |
| 11 | 130.4 | 4.7 | 2095 | 22 | AAE44168 Human PRO1074 (UNQ |
| 12 | 130.4 | 4.7 | 2168 | 22 | AAH15711 Human cDNA sequenc |
| 13 | 130.4 | 4.7 | 2189 | 20 | AAH87192 Human Dendriac cDN |
| 14 | 124.8 | 4.5 | 333 | 20 | AAH51967 Human secreted pro |
| 15 | 124 | 4.5 | 2420 | 20 | AAH35711 cDNA encoding a pr |
| 16 | 122.4 | 4.4 | 1266 | 20 | AAH35710 cDNA encoding a pr |
| 17 | 119.8 | 4.3 | 847 | 21 | AAH69087 Human secreted pro |
| 18 | 115.8 | 4.2 | 9289 | 22 | AAH30960 Human digestive sy |
| 19 | 115.8 | 4.2 | 9289 | 22 | AAH31995 Human liver associ |
| 20 | 115.8 | 4.2 | 9289 | 24 | ABN90350 Human liver antige |
| 21 | 115.8 | 4.2 | 17865 | 22 | AAH90961 Human digestive sy |
| 22 | 115.8 | 4.2 | 17865 | 22 | AAH31996 Human liver associ |
| 23 | 115.8 | 4.2 | 17865 | 24 | ABN90351 Human liver antige |
| 24 | 114.8 | 4.1 | 1134 | 21 | AAH58791 DNA encoding the b |
| 25 | 111.6 | 4.0 | 671 | 22 | AAH52784 Human polynucleoti |
| 26 | 111.6 | 4.0 | 827 | 22 | AAH51800 Human polynucleoti |
| 27 | 108.4 | 3.9 | 125439 | 24 | ABO88177 Human osteoblast d |
| 28 | 108.2 | 3.9 | 1134 | 22 | AAH09356 Human beta-1,3-gal |
| 29 | 106 | 3.8 | 303 | 24 | ABN24320 Human ORFX polynuc |
| 30 | 105 | 3.8 | 474 | 22 | ABA59003 Human foetal liver |
| 31 | 105 | 3.8 | 474 | 22 | ABA27841 Probe #6307 for ge |
| 32 | 105 | 3.8 | 474 | 22 | AAK07173 Human brain expres |
| 33 | 105 | 3.8 | 474 | 22 | AAK32922 Human bone marrow |
| 34 | 105 | 3.8 | 474 | 22 | AAI16260 Probe #6193 for ge |
| 35 | 105 | 3.8 | 474 | 22 | AAI38728 Probe #7414 used t |
| 36 | 105 | 3.8 | 474 | 24 | ABS07727 Human genome-deriv |
| 37 | 103.8 | 3.7 | 37808 | 20 | AAH02780 Vector PMVX-BG DNA |
| 38 | 97.8 | 3.5 | 160271 | 22 | AAF85750 Bipolar affective |
| 39 | 97.8 | 3.5 | 160271 | 22 | AAF85756 Human chromosome 1 |
| 40 | 97.8 | 3.5 | 160271 | 22 | AAH04858 Human chromosome 1 |
| 41 | 97.8 | 3.5 | 160271 | 22 | AAH04864 Human chromosome 1 |
| 42 | 97.8 | 3.5 | 160271 | 22 | AAH06667 Human chromosome 1 |
| 43 | 97.8 | 3.5 | 160271 | 22 | AAH40997 160kb fragment of |
| 44 | 97.8 | 3.5 | 160271 | 22 | AAH23764 Human chromosome 1 |
| 45 | 97.8 | 3.5 | 160271 | 22 | AAF85116 Nucleotide sequenc |

ALIGNMENTS

RESULT 1
AAA93875
ID AAA93875 standard; DNA; 2775 BP.

XX AAA93875;

XX AC AAA93875;

XX 15-JAN-2001 (first entry)

XX Human beta-1,3 galactose transferase encoding DNA.

XX Beta-1,3 galactose transferase; treatment; diagnosis; cancer; human;

XX digestive system; ds.

XX Homo sapiens.

XX WO2000050608-A1.

XX 31-AUG-2000.

XX 24-FEB-2000; 2000WO-JP01070.

XX 25-FEB-1999; 99JP-0047571.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Narimatsu H, Isshiki S, Togayachi A, Sasaki K;

XX WPI: 2000-549409/50.

XX P-PSDB; AAB93875.

XX Beta-1,3 galactose transferase and DNA encoding it, useful for

XX synthesis of type 1 sialyl Lewis, a carbohydrate for treatment of

PT digestive system cancer -
XX
PS Claim 5; Page 99-102; 123pp; Japanese.
XX
CC This invention relates to a polypeptide (I) with beta-1,3 galactose
CC transferase activity, or variants of (I) comprising amino acid additions,
CC deletions and/or substitutions. Included in the invention is DNA encoding
CC all or part of (I); expression vectors containing the DNA, host cells
CC transformed by the vectors; a method for the preparation of the
CC polypeptide by culture of the transformants or by expression in the milk
CC of a transgenic mammal, and antibodies recognising (I). The Beta-1,3
CC galactose transferase protein transfers galactose by beta-1,3 bonding to
CC N-acetylglucosamine present in a non-cyclic carbohydrate chain (such as
CC GlcNAc) to give Galbeta1-3GlcNAc. The protein and
CC DNA encoding it are useful for the treatment and diagnosis of cancer of
CC the digestive system. The present sequence represents Beta-1,3 galactose
CC transferase encoding DNA.
XX
SQ Sequence 2775 BP; 681 A; 698 G; 669 G; 727 T; 0 other;

Query Match 100.0%; Score 2775; DB 21; Length 2775;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2775; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGAATTCCTCTTCTCTGCTGGAGCTGGGATATCTTCTCTCTCCCTGGACATCA 60
DB 1 GTGAATTCCTCTTCTCTGCTGGAGCTGGGATATCTTCTCTCTCCCTGGACATCA 60
QY 61 GAGCTGCAGGCTCTCTGGCCCTTTGGACCCGAGGATTTATACCAAGCAGGTTTCTGGGTT 120
DB 61 GAGCTGCAGGCTCTCTGGCCCTTTGGACCCGAGGATTTATACCAAGCAGGTTTCTGGGTT 120
QY 121 TCAGGCTTTGGCCCTTGGAGCTAGTATACACCATGGCATATCTGGTCTGAGGCTCTT 180
DB 121 TCAGGCTTTGGCCCTTGGAGCTAGTATACACCATGGCATATCTGGTCTGAGGCTCTT 180
QY 181 GGTCTTGGACTGAGCCACTCTCTGGCATCCGAGCTCTCCAGCTTCGATGGCTGTAC 240
DB 181 GGTCTTGGACTGAGCCACTCTCTGGCATCCGAGCTCTCCAGCTTCGATGGCTGTAC 240
QY 241 GTGATTCCTGTGAGATCAGCATTTTGGTAAACAAACGAGCCGAGCATGATAATTA 300
DB 241 GTGATTCCTGTGAGATCAGCATTTTGGTAAACAAACGAGCCGAGCATGATAATTA 300
QY 301 TGGAGCATTTTACACTGACAGTCTTTGAGACAAATTTCTCTTGGCATTTACACTGTGG 360
DB 301 TGGAGCATTTTACACTGACAGTCTTTGAGACAAATTTCTCTTGGCATTTACACTGTGG 360
QY 361 CTTTGTAGCTTTCAAACAGAGGTTCTCTTACCCAGCAAAATGGCTTTCCCGAAGATGA 420
DB 361 CTTTGTAGCTTTCAAACAGAGGTTCTCTTACCCAGCAAAATGGCTTTCCCGAAGATGA 420
QY 421 GATTGATGTATATTGGCTTCTGGTCTGGGGGCTCTTTGTTGTATTTAGCATGTACA 480
DB 421 GATTGATGTATATTGGCTTCTGGTCTGGGGGCTCTTTGTTGTATTTAGCATGTACA 480
QY 481 GTCTAAATCTTTCAAAGACAGTCTTTGTTTACAGAAAGAGCGGAATCTTCCCTTAAGC 540
DB 481 GTCTAAATCTTTCAAAGACAGTCTTTGTTTACAGAAAGAGCGGAATCTTCCCTTAAGC 540
QY 541 TCCAGATACAGACTGCAGGAGACACTCCCTTCTCTGCTGCTGGTGAACCTCATCCC 600
DB 541 TCCAGATACAGACTGCAGGAGACACTCCCTTCTCTGCTGCTGGTGAACCTCATCCC 600
QY 601 ACAACAGTTGGCTGAGCGCATGCCATCCGAGAGCTGGGGAAAGAGAGGATGTGA 660
DB 601 ACAACAGTTGGCTGAGCGCATGCCATCCGAGAGCTGGGGAAAGAGAGGATGTGA 660
QY 661 AGGAAAGCAGCTGAAGACATTTCTTCTCTGGGAGCCACAGCAGTGCAGCGGAACGA 720
DB 661 AGGAAAGCAGCTGAAGACATTTCTTCTCTGGGAGCCACAGCAGTGCAGCGGAACGA 720
QY 721 AAGAGTGGACAGGAGGAGCCAGGACACGGGACATTTATCCAGAGGATTTCTCTAGACG 780

DB 721 AAGAGTGGACAGGAGGAGCCAGGACACGGGACATTTATCCAGAGGATTTCTCTAGACG 780
QY 781 TCTATTACAATCTGACCTGAAGACCATGATGGGATAGAAATGGGTCCTGCTTTGTC 840
DB 781 TCTATTACAATCTGACCTGAAGACCATGATGGGATAGAAATGGGTCCTGCTTTGTC 840
QY 841 CTCAGGGGGCTTTGTGTGATGAAACAGACTCAGACATGTTCAATGTTGACTATCTGA 900
DB 841 CTCAGGGGGCTTTGTGTGATGAAACAGACTCAGACATGTTCAATGTTGACTATCTGA 900
QY 901 CTGAACCTCTTCTGAAGAAACAGAAACACAGGTTTTTCTGCTGCTTTTGAACATCA 960
DB 901 CTGAACCTCTTCTGAAGAAACAGAAACACAGGTTTTTCTGCTGCTTTTGAACATCA 960
QY 961 ATGAGTTTCCCATCAGGCGAGCCATTTACGAAAGTGGTTGTGAGTAAATCTGAATATCCGT 1020
DB 961 ATGAGTTTCCCATCAGGCGAGCCATTTACGAAAGTGGTTGTGAGTAAATCTGAATATCCGT 1020
QY 1021 GGGACAGGTACCCACCATTTCTGCTCCGCGACCGGCTAGCTGTTTCTGGCGACGTGGGA 1080
DB 1021 GGGACAGGTACCCACCATTTCTGCTCCGCGACCGGCTAGCTGTTTCTGGCGACGTGGGA 1080
QY 1081 GTCAGGTGTACAATGTCCTCAAGAGCGTCCCATACATTTAAACTGGAAGACGTTTGTGG 1140
DB 1081 GTCAGGTGTACAATGTCCTCAAGAGCGTCCCATACATTTAAACTGGAAGACGTTTGTGG 1140
QY 1141 GGCTCTGCTCCAAAGGCTGAACATCAGATTTGGAGAGCTCCACTCCCGACGACCTTTT 1200
DB 1141 GGCTCTGCTCCAAAGGCTGAACATCAGATTTGGAGAGCTCCACTCCCGACGACCTTTT 1200
QY 1201 TTCAGGGGGCTTACGCTTCTCCGATGCCCTTTCAGGAGGATCGTGCCCTGCCACTTCA 1260
DB 1201 TTCAGGGGGCTTACGCTTCTCCGATGCCCTTTCAGGAGGATCGTGCCCTGCCACTTCA 1260
QY 1261 TCAAGCTCTGAGCTCTCTTGGAGTCTAGAGAAATTTCCCGGGGGAGGATTT 1320
DB 1261 TCAAGCTCTGAGCTCTCTTGGAGTCTAGAGAAATTTCCCGGGGGAGGATTT 1320
QY 1321 GTCGCGCTGTCTGAGGGGAGCCAGAGGACATCCGGAAGATTTTCAGATAACCGTGGG 1380
DB 1321 GTCGCGCTGTCTGAGGGGAGCCAGAGGACATCCGGAAGATTTTCAGATAACCGTGGG 1380
QY 1381 GATAGTTTTCGCTAGATTTTGGAGAGGGGGGAGAGAGATGCTGCTTCTAGTCT 1440
DB 1381 GATAGTTTTCGCTAGATTTTGGAGAGGGGGGAGAGAGATGCTGCTTCTAGTCT 1440
QY 1441 GAAATCCACGCGCAGATGTCGGTGTTCATGAAGTCACTGATTAGTCCCACTTGGTGCC 1500
DB 1441 GAAATCCACGCGCAGATGTCGGTGTTCATGAAGTCACTGATTAGTCCCACTTGGTGCC 1500
QY 1501 CAGGCAATATAGCCCGCTCTCTTGGGACGACACACTTTCATATAAGTGTGTTGACATA 1560
DB 1501 CAGGCAATATAGCCCGCTCTCTTGGGACGACACACTTTCATATAAGTGTGTTGACATA 1560
QY 1561 CACCTGATTTTTCATTTTTCAGGGGTGATGATGATGATGATGATGATGATGATGATGAT 1620
DB 1561 CACCTGATTTTTCATTTTTCAGGGGTGATGATGATGATGATGATGATGATGATGATGAT 1620
QY 1621 AATTTTACAGGCAAGGACACAGAGCTGCGAGAGGTGAGAGGATGTTGCCAAGGCTCAC 1680
DB 1621 AATTTTACAGGCAAGGACACAGAGCTGCGAGAGGTGAGAGGATGTTGCCAAGGCTCAC 1680
QY 1681 AGCCAGTAGGCATAGGAGCGGGAATGAAATTCGAGACTGTCAGAAATCTGTTGGGACGCC 1740
DB 1681 AGCCAGTAGGCATAGGAGCGGGAATGAAATTCGAGACTGTCAGAAATCTGTTGGGACGCC 1740
QY 1741 CCTGACTTGAACCACTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
DB 1741 CCTGACTTGAACCACTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
QY 1801 TCGAGCGCGGCTCTTCTCCATCCCTGTCGAGTCCCTCCACTCAGCTCCCGCTGCTTGT 1860

Db 1801 TCGAGCGCGCATCTCCATCCCTGTCGAGTCCCTCCACCTCAGCTCCCGCTCTGT 1860
QY 1861 GCTTTTGGAGCTAAGCTGGGATGACCAAAATTCACCCAGCTCCTTCATTCACAGGGCT 1920
Db 1861 GCTTTTGGAGCTAAGCTGGGATGACCAAAATTCACCCAGCTCCTTCATTCACAGGGCT 1920
QY 1921 GGATGAGCTGGGATGAGTCCATGTTATCGGCTCGGTACTCAACACACCAAGTTTCA 1980
Db 1921 GGATGAGCTGGGATGAGTCCATGTTATCGGCTCGGTACTCAACACACCAAGTTTCA 1980
QY 1981 TCCGAGGAAATGCTCCCGCAGTGGATGAGCTCAGATGCTGAGAACACCCAGCTCTGA 2040
Db 1981 TCCGAGGAAATGCTCCCGCAGTGGATGAGCTCAGATGCTGAGAACACCCAGCTCTGA 2040
QY 2041 CAGAGTCTCTTATAAATGTATAAATAGGCTCAGAAACCACTGCATTTCTGACCTGTGTAC 2100
Db 2041 CAGAGTCTCTTATAAATGTATAAATAGGCTCAGAAACCACTGCATTTCTGACCTGTGTAC 2100
QY 2101 AGACTGCCACACTGCTGACCTGCTAGGAGCAGACATCCCTTCTGAGCCATCTGTG 2160
Db 2101 AGACTGCCACACTGCTGACCTGCTAGGAGCAGACATCCCTTCTGAGCCATCTGTG 2160
QY 2161 CTTCTCTCATTTTCATCACCCTCACTGCTCCCTGCTTTTGTATCAATGGGACGACCTGC 2220
Db 2161 CTTCTCTCATTTTCATCACCCTCACTGCTCCCTGCTTTTGTATCAATGGGACGACCTGC 2220
QY 2221 CCCAGGAGCATTAGGGCTCTCAGTTCAACTGAAGACAGCTTGAACCTCAGATGGGGTT 2280
Db 2221 CCCAGGAGCATTAGGGCTCTCAGTTCAACTGAAGACAGCTTGAACCTCAGATGGGGTT 2280
QY 2281 CATGTGGGATTCGGGAGCTTCTGGGAATTCAGTTGGAGTCAAGTGCCTCAA 2340
Db 2281 CATGTGGGATTCGGGAGCTTCTGGGAATTCAGTTGGAGTCAAGTGCCTCAA 2340
QY 2341 GGACCCCTCGGCTCAGAGCCCTAAGTGGGCTTGAAGCAGGGTGGTCTCGCTCC 2400
Db 2341 GGACCCCTCGGCTCAGAGCCCTAAGTGGGCTTGAAGCAGGGTGGTCTCGCTCC 2400
QY 2401 ACTTCCCAAGCTGAGCCAAAGCTATCTTCAATGATGCTCATTTGAATGCTCATTTGGCCGAGGAACAAC 2460
Db 2401 ACTTCCCAAGCTGAGCCAAAGCTATCTTCAATGATGCTCATTTGAATGCTCATTTGGCCGAGGAACAAC 2460
QY 2461 TGAACCTTTGTGGTTTGTGCTTTAGCTTCACTTCTGCTCCGCTGCTTACCCAGAGTT 2520
Db 2461 TGAACCTTTGTGGTTTGTGCTTTAGCTTCACTTCTGCTCCGCTGCTTACCCAGAGTT 2520
QY 2521 TGTGCGAGCTGCTGTGCGAGGTTGTATAAAACCAAGTACTTCTGTTAGTTTGGCCATT 2580
Db 2521 TGTGCGAGCTGCTGTGCGAGGTTGTATAAAACCAAGTACTTCTGTTAGTTTGGCCATT 2580
QY 2581 CAGCCATGGTCACGTGACATGCAAAAGTAACTTGTCTCTTAATATAGAAATGATTTTCT 2640
Db 2581 CAGCCATGGTCACGTGACATGCAAAAGTAACTTGTCTCTTAATATAGAAATGATTTTCT 2640
QY 2641 TTTAATTTTACTTTTACCAGCTTTACTTGTACTCAGAGAGAGGCTCAGATGGCTG 2700
Db 2641 TTTAATTTTACTTTTACCAGCTTTACTTGTACTCAGAGAGAGGCTCAGATGGCTG 2700
QY 2701 TGTACATATAAATGTTGAGCTTAACTCTTAAAAAATAAATAAATAAATAAATAAATAA 2760
Db 2701 TGTACATATAAATGTTGAGCTTAACTCTTAAAAAATAAATAAATAAATAAATAAATAA 2760
QY 2761 AAAAAAAAAAAAAA 2775
Db 2761 AAAAAAAAAAAAAA 2775

RESULT 2
ID AAA93876
XX AAA93876 standard; DNA; 10562 BP.
AC
XX AAA93876;
XX

DT 15-JAN-2001 (first entry)
XX Human beta3Gal-T5 encoding DNA.
DE
XX Beta-1,3 galactose transferase; treatment; diagnosis; cancer; human;
KW digestive system; beta3Gal-T5; ds.
KW Homo sapiens.
OS
XX WO2000050608-A1.
PN
XX 31-AUG-2000.
PD
XX 24-FEB-2000; 2000WO-JP01070.
PF
XX 25-FEB-1999; 99JP-0047571.
PR
XX (KYOM) KYOMA HAKKO KOGYO KK.
PA
XX Narimatsu H, Isshiki S, Togayachi A, Sasaki K;
PI
XX WPI; 2000-549409/50.
DR
XX Beta-1,3 galactose transferase and DNA encoding it, useful for
PT synthesis of type 1 sialyl Lewis^x, a carbohydrate for treatment of
PT digestive system cancer
PS
XX Claim 31; Page 103-111; 123pp; Japanese.
XX This invention relates to a polypeptide (I) with beta-1,3 galactose
CC transferase activity, or variants of (I) comprising amino acid additions,
CC deletions and/or substitutions. Included in the invention is DNA encoding
CC all or part of (I); expression vectors containing the DNA, host cells
CC transformed by the vectors; a method for the preparation of the
CC polypeptide by culture of the transformants or by expression in the milk
CC of a transgenic mammal, and antibodies recognising (I). The beta-1,3
CC galactose transferase protein transfers galactose by beta-1,3 bonding to
CC N-acetylglucosamine present in a non-cyclic carbohydrate chain (such as
CC GlcNAc-beta1-3Gal-beta1-4Glc) to give Gal-beta1-3GlcNAc. The protein and
CC DNA encoding it are useful for the treatment and diagnosis of cancer of
CC the digestive system. The present sequence represents a Beta3gal-T5
CC encoding DNA sequence.
XX
SQ Sequence 10562 BP; 2610 A; 2415 C; 2574 G; 2963 T; 0 other;

Query Match 83.9%; Score 2329.4; DB 21; Length 10562;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2330; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 400 AAATGGCTTTCCCGAAGATGAGATTGATGATATATTTGCCCTTCTGGTTCTGGGGCTCTTT 459
Db 8232 AGATGGCTTTCCCGAAGATGAGATTGATGATATATTTGCCCTTCTGGTTCTGGGGCTCTTT 8291
QY 460 GTTTGTATTTAGCATGTACAGTCTAAATTCCTTTCAAAGAACAGTCTCTTTGTTTACAAGA 519
Db 8292 GTTTGTATTTAGCATGTACAGTCTAAATTCCTTTCAAAGAACAGTCTCTTTGTTTACAAGA 8351
QY 520 AAGCGGGGAACCTTCCCTTAAGCTCCCGATACAGCTGCGAGGAGACACCTCCCTCTCG 579
Db 8352 AAGCGGGGAACCTTCCCTTAAGCTCCCGATACAGCTGCGAGGAGACACCTCCCTCTCG 8411
QY 580 TCCTGCTGTGACCTCATCCCAACACAGTTGGCTGAGCGCATGGCGATCCGCGACACCT 639
Db 8412 TCCTGCTGTGACCTCATCCCAACACAGTTGGCTGAGCGCATGGCGATCCGCGACACCT 8471
QY 640 GGGGAAAAGAGAGGATGGTGAAGGAAAAGACAGCTGAAGACATTTCTTCTCTCTGGGACCA 699
Db 8472 GGGGAAAAGAGAGGATGGTGAAGGAAAAGACAGCTGAAGACATTTCTTCTCTCTGGGACCA 8531
QY 700 CCAGCAGTGCACCGGAAACGAAGAGGTGGACACAGGAGGAGGAGGAGGACATTA 759
Db 8532 CCAGCAGTGCACCGGAAACGAAGAGGTGGACACAGGAGGAGGAGGAGGACATTA 8591

| | | | |
|----|------|--|------|
| Qy | 760 | TCCAGAAGGATTTCTCTAGACGCTATTAACTATGACCCCTGAAGACCATGATGGCATAG | 819 |
| Db | 8592 | | |
| | | TCCAGAAGGATTTCTCTAGACGCTATTAACTATGACCCCTGAAGACCATGATGGCATAG | 8651 |
| Qy | 820 | AATGGGCTCCATCGCTTTTGTCTCTCAGCGCGGTTTGTGATGAAACAGACTCAGACATGT | 879 |
| Db | 8652 | | |
| | | AATGGGCTCCATCGCTTTTGTCTCTCAGCGCGGTTTGTGATGAAACAGACTCAGACATGT | 8711 |
| Qy | 880 | TCATCAATGTTGACTATCTGACTGAACCTGCTTCTTGAAGAAAAACAGAACACAGGTTTTT | 939 |
| Db | 8712 | | |
| | | TCATCAATGTTGACTATCTGACTGAACCTGCTTCTTGAAGAAAAACAGAACACAGGTTTTT | 8771 |
| Qy | 940 | TCACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGCGACCCATTTCAGCAAGTGGTTTG | 999 |
| Db | 8772 | | |
| | | TCACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGCGACCCATTTCAGCAAGTGGTTTG | 8831 |
| Qy | 1000 | TCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCAATTCGTCTCGGCGACCGGCTACG | 1059 |
| Db | 8832 | | |
| | | TCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCAATTCGTCTCGGCGACCGGCTACG | 889 |
| Qy | 1060 | TGTTTTCTGGCGACGTGGCGAGTCAAGTGTACAATGTCTCCAAGAGCGTCCCATACATTA | 1119 |
| Db | 8892 | | |
| | | TGTTTTCTGGCGACGTGGCGAGTCAAGTGTACAATGTCTCCAAGAGCGTCCCATACATTA | 8951 |
| Qy | 1120 | AAC TGAAGACGTTGTTGTGGGCTCTGCCTCGAAGGCTGAACATCAGATTGGAGGAGC | 1179 |
| Db | 8952 | | |
| | | AAC TGAAGACGTTGTTGTGGGCTCTGCCTCGAAGGCTGAACATCAGATTGGAGGAGC | 9011 |
| Qy | 1180 | TCACTCCGACGCGACCTTTTTTCCAGGGGGTTACGGTCTCCGTATGCGCTCTTCAGGA | 1239 |
| Db | 9012 | | |
| | | TCCACTCCGACGCGACCTTTTTTCCAGGGGGTTACGGTCTCCGTATGCGCTCTTCAGGA | 9071 |
| Qy | 1240 | GGATCGTGGCTGCCACTTCATCAAGCGCTCGGACTCTCTTGGACTACTGCGAGGCTCTAG | 1299 |
| Db | 9072 | | |
| | | GGATCGTGGCTGCCACTTCATCAAGCGCTCGGACTCTCTTGGACTACTGCGAGGCTCTAG | 9131 |
| Qy | 1300 | AGAA TCCCGGGGGGAAGATTGTCGCGCTGTCTGAGGGGAGCCCGAGGCGACATPCGGGAC | 1359 |
| Db | 9132 | | |
| | | AGAA TCCCGGGGGGAAGATTGTCGCGCTGTCTGAGGGGAGCCCGAGGCGACATPCGGGAC | 9191 |
| Qy | 1360 | AAGTTTCAGATAACCCGCTGGGATAGTTTTTGTCTAGATTTTGAAGAGGGGGCGGACAG | 1419 |
| Db | 9192 | | |
| | | AAGTTTCAGATAACCCGCTGGGATAGTTTTTGTCTAGATTTTGAAGAGGGGGCGGACAG | 9251 |
| Qy | 1420 | AGGATGCTGTTCTTCACTGCTGAAATCCACGCCAGATGTCGGTGTTCATGAAGTCACTG | 1479 |
| Db | 9252 | | |
| | | AGGATGCTGTTCTTCACTGCTGAAATCCACGCCAGATGTCGGTGTTCATGAAGTCACTG | 9311 |
| Qy | 1480 | ATTAGTTCCCACTTGTGTCGCCCAGGCAATAATAGGCCCGTCTCTTGGGACGACACTCT | 1539 |
| Db | 9312 | | |
| | | ATTAGTTCCCACTTGTGTCGCCCAGGCAATAATAGGCCCGTCTCTTGGGACGACACTCT | 9371 |
| Qy | 1540 | TCATACTAAGTGTTCGACATACACCTGGATTTTTCATTTCCAGGGGTCAAGTATCTATGA | 1599 |
| Db | 9372 | | |
| | | TCATACTAAGTGTTCGACATACACCTGGATTTTTCATTTCCAGGGGTCAAGTATCTATGA | 9431 |
| Qy | 1600 | CATGATGGGTGTTTACCATCCTAATTTTACAGCGAGGACACAGCAGCTGGGAGAGGTACA | 1659 |
| Db | 9432 | | |
| | | CATGATGGGTGTTTACCATCCTAATTTTACAGCGAAGGACACAGCAGCTGGGAGAGGTACA | 9491 |
| Qy | 1660 | GAAACTTGTCCCAAGGCTCACAGCCAGTAGGCGATAGGAGCGGGGAATGAAAATCGAGCACT | 1719 |
| Db | 9492 | | |
| | | GAAACTTGTCCCAAGGCTCACAGCCAGTAGGCGATAGGAGCGGGGAATGAAAATCGAGCACT | 9551 |
| Qy | 1720 | GTCAAGATCTGTGGGACGCCCTGACTTGAACCACTCCGACGTGCTGCCCTCCCTTAGGA | 1779 |
| Db | 9552 | | |
| | | GTCAAGATCTGTGGGACGCCCTGACTTGAACCACTCCGACGTGCTGCCCTCCCTTAGGA | 9611 |
| Qy | 1780 | GGGGACACTCATGATGAGTCTCGGAGCGCGCATCCTTCCATCCCTGTGAGTCCCGCTCC | 1839 |
| Db | 9612 | | |
| | | GGGGACACTCATGATGAGTCTCGGAGCGCGCATCCTTCCATCCCTGTGAGTCCCGCTCC | 9671 |
| Qy | 1840 | ACCTCAGCTCCCGACTCTTGTGCTTTTGGAGCTAAGCGCTGGGATGACCAAAATTCACCCC | 1899 |

| | | | | |
|----|-------|--|---|-------|
| Db | 9672 | | ACCTCAGCTCCAGTCCCTTGTCGCTTTTGGAGCTTAGGCTGGATGACCAAAATTCACCC | 9731 |
| Qy | 1900 | | AGCTCTTTCATTACAGGGCTGGATGAGCTGGGATTTAGTCCATGTTATCGGCTCGGTA | 1959 |
| Db | 9732 | | AGCTCCTTCATTACAGGGCTGGATGAGCTGGGATTTAGTCCATGTTATCGGCTCGGTA | 9791 |
| Qy | 1960 | | CTCAACAACCCCAAGTTTCATCCGAGAAATGTCGCCGAGTGAGTGCAGCTCACATGC | 2019 |
| Db | 9792 | | CTCAACAACCCCAAGTTTCATCCGAGAAATGTCGCCGAGTGAGTGCAGCTCACATGC | 9851 |
| Qy | 2020 | | TGAGGAACACCCAGCTCTGGACAGAGTTCTTTATAAATGATATAAATAGGCTCGAAAACCA | 2079 |
| Db | 9852 | | TGAGGAACACCCAGCTCTGGACAGAGTTCTTTATAAATGATATAAATAGGCTCGAAAACCA | 9911 |
| Qy | 2080 | | CTGCATTTCTGACCTGCTACAGAGTGCACACTGCTGACCTGCCTAGCAGCAGGACA | 2139 |
| Db | 9912 | | CTGCATTTCTGACCTGCTACAGAGTGCACACTGCTGACCTGCCTAGCAGCAGGACA | 9971 |
| Qy | 2140 | | TCCCTTCTGAGCCATCTGCTCTCTCATTTTCATCACCCCAACTGTCCTTGTGTTTGA | 2199 |
| Db | 9972 | | TCCCTTCTGAGCCATCTGCTCTCTCATTTTCATCACCCCAACTGTCCTTGTGTTTGA | 10031 |
| Qy | 2200 | | TCAATGGGACACGCACTGCCCCAGGAGCACTTTAGGCTCTCAGTTTCAAACTGAAGGA | 2259 |
| Db | 10032 | | TCAATGGGACACGCACTGCCCCAGGAGCACTTTAGGCTCTCAGTTTCAAACTGAAGGA | 10091 |
| Qy | 2260 | | CAGTTTGAATCAGATGGGGTTTCATGTGGATTCTGGGAGCTTCTTGGGAAATTCAGTTGA | 2319 |
| Db | 10092 | | CAGTTTGAATCAGATGGGGTTTCATGTGGATTCTGGGAGCTTCTTGGGAAATTCAGTTGA | 10151 |
| Qy | 2320 | | GTCAAGTCAGGATGCTCTCAAGGACCCCTCGGGCTCAGAGCCCTAAAGTGGGCCCTGGTG | 2379 |
| Db | 10152 | | GTCAAGTCAGGATGCTCTCAAGGACCCCTCGGGCTCAGAGCCCTAAAGTGGGCCCTGGTG | 10211 |
| Qy | 2380 | | AAGCAGGTTGTCCTGCGTCCACTTCCCAAGCTTGAGCCAAGCTCATCTTCATTGAATGT | 2439 |
| Db | 10212 | | AAGCAGGTTGTCCTGCGTCCACTTCCCAAGCTTGAGCCAAGCTCATCTTCATTGAATGT | 10271 |
| Qy | 2440 | | CTCATTTGGCCGAGGAACAACCTGAACCTTGTTGGTTTGTCTTAGCCTTCAGTTTGTCTC | 2499 |
| Db | 10272 | | CTCATTTGGCCGAGGAACAACCTGAACCTTGTTGGTTTGTCTTAGCCTTCAGTTTGTCTC | 10331 |
| Qy | 2500 | | GCTGCCTCTACCCAGAGTTTGTGCGAGCCTGTGTCAGGGTTGTATAAAACCAAGGT | 2559 |
| Db | 10332 | | GCTGCCTCTACCCAGAGTTTGTGCGAGCCTGTGTCAGGGTTGTATAAAACCAAGGT | 10391 |
| Qy | 2560 | | ACTTCGTTAGTTTTCGCCATTCAGCCATGGTCACGTGCATGCAAGTAATCTTGTCTCCT | 2619 |
| Db | 10392 | | ACTTCGTTAGTTTTCGCCATTCAGCCATGGTCACGTGCATGCAAGTAATCTTGTCTCCT | 10451 |
| Qy | 2620 | | AATTATAGAAATGATTTTCTTTTAAATTTTACTTTTACCAGACTTACTTTGTACTCAG | 2679 |
| Db | 10452 | | AATTATAGAAATGATTTTCTTTTAAATTTTACTTTTACCAGACTTACTTTGTACTCAG | 10511 |
| Qy | 2680 | | AGAAGAGGCTCACATGGCTGTGCACATATAAATGTTGGACTAAACTCTT | 2730 |
| Db | 10512 | | ACAAGAGCTCACATGGCTGTGCACATATAAATGTTGGACTAAACTCTT | 10562 |

RESULT 3
AAA27959

AAA27959
ID AAA27959 standard; DNA: 1011 BP.

AAA
AC AAA27959;

DT 15-AUG-2000 (first entry)

XX Human beta3Gal-T5 gene sequence. DE

XX
KW UDP-D-galactose; beta-N-acetylglucosamine beta1,3-galactosyltransferase;
KW Beta3Gal-T5; ss; human; chromosome 21q22.3; galactosylation;
KW beta1,3-galactosyl glycosylated saccharide production; glycopeptide;

XX Human beta3Gal-T5 gene sequence. DE

461 TAGGCCAAGAGGCTGAAAGGAAGACAAATGTTGGCATGTCTTAGAGGATGAACACC 520

04-JUN-1998: 98US-0088033
04-JUN-1998: 98US-0088030
PR

PR 04-JUN-1998; 98US-0088326.
PR 05-JUN-1998; 98US-0088167.
PR 05-JUN-1998; 98US-0088202.
PR 05-JUN-1998; 98US-0088212.
PR 05-JUN-1998; 98US-0088217.
PR 09-JUN-1998; 98US-0088655.
PR 10-JUN-1998; 98US-0088722.
PR 10-JUN-1998; 98US-0088730.
PR 10-JUN-1998; 98US-0088734.
PR 10-JUN-1998; 98US-0088738.
PR 10-JUN-1998; 98US-0088740.
PR 10-JUN-1998; 98US-0088741.
PR 10-JUN-1998; 98US-0088742.
PR 10-JUN-1998; 98US-0088810.
PR 10-JUN-1998; 98US-0088811.
PR 10-JUN-1998; 98US-0088824.
PR 10-JUN-1998; 98US-0088825.
PR 10-JUN-1998; 98US-0088826.
PR 11-JUN-1998; 98US-0088858.
PR 11-JUN-1998; 98US-0088861.
PR 11-JUN-1998; 98US-0088863.
PR 11-JUN-1998; 98US-0088876.
PR 12-JUN-1998; 98US-0089090.
PR 12-JUN-1998; 98US-0089105.
PR 16-JUN-1998; 98US-0089440.
PR 16-JUN-1998; 98US-0089512.
PR 16-JUN-1998; 98US-0089514.
PR 17-JUN-1998; 98US-0089532.
PR 17-JUN-1998; 98US-0089538.
PR 17-JUN-1998; 98US-0089598.
PR 17-JUN-1998; 98US-0089599.
PR 17-JUN-1998; 98US-0089600.
PR 17-JUN-1998; 98US-0089653.
PR 18-JUN-1998; 98US-0089801.
PR 18-JUN-1998; 98US-0089907.
PR 18-JUN-1998; 98US-0089908.
PR 19-JUN-1998; 98US-0089947.
PR 19-JUN-1998; 98US-0089948.
PR 19-JUN-1998; 98US-0089952.
PR 22-JUN-1998; 98US-0090246.
PR 22-JUN-1998; 98US-0090252.
PR 22-JUN-1998; 98US-0090254.
PR 23-JUN-1998; 98US-0090349.
PR 23-JUN-1998; 98US-0090355.
PR 24-JUN-1998; 98US-0090429.
PR 24-JUN-1998; 98US-0090431.
PR 24-JUN-1998; 98US-0090435.
PR 24-JUN-1998; 98US-0090444.
PR 24-JUN-1998; 98US-0090445.
PR 24-JUN-1998; 98US-0090461.
PR 24-JUN-1998; 98US-0090472.
PR 24-JUN-1998; 98US-0090535.
PR 24-JUN-1998; 98US-0090538.
PR 24-JUN-1998; 98US-0090540.
PR 24-JUN-1998; 98US-0090557.
PR 25-JUN-1998; 98US-0090676.
PR 25-JUN-1998; 98US-0090678.
PR 25-JUN-1998; 98US-0090688.
PR 25-JUN-1998; 98US-0090690.
PR 25-JUN-1998; 98US-0090691.
PR 25-JUN-1998; 98US-0090694.
PR 25-JUN-1998; 98US-0090695.
PR 25-JUN-1998; 98US-0090696.
PR 26-JUN-1998; 98US-0090696.
PR 26-JUN-1998; 98US-0090863.
PR 01-JUL-1998; 98US-0091358.
PR 01-JUL-1998; 98US-0091360.
PR 01-JUL-1998; 98US-0091544.
PR 02-JUL-1998; 98US-0091478.
PR 02-JUL-1998; 98US-0091486.
PR 02-JUL-1998; 98US-0091519.
PR 02-JUL-1998; 98US-0091626.
PR 02-JUL-1998; 98US-0091628.
PR 02-JUL-1998; 98US-0091633.
PR 02-JUL-1998; 98US-0091646.
PR 02-JUL-1998; 98US-0091673.
PR 07-JUL-1998; 98US-0091978.
PR 07-JUL-1998; 98US-0091982.
PR 09-JUL-1998; 98US-0092182.
PR 10-JUL-1998; 98US-0092472.
PR 20-JUL-1998; 98US-0093339.
PR 30-JUL-1998; 98US-0094651.
PR 04-AUG-1998; 98US-0095282.
PR 04-AUG-1998; 98US-0095285.
PR 04-AUG-1998; 98US-0095301.
PR 04-AUG-1998; 98US-0095302.
PR 04-AUG-1998; 98US-0095318.
PR 04-AUG-1998; 98US-0095321.
PR 04-AUG-1998; 98US-0095325.
PR 10-AUG-1998; 98US-0095916.
PR 10-AUG-1998; 98US-0096012.
PR 11-AUG-1998; 98US-0096013.
PR 11-AUG-1998; 98US-0096146.
PR 12-AUG-1998; 98US-0096329.
PR 17-AUG-1998; 98US-0096757.
PR 17-AUG-1998; 98US-0096766.
PR 17-AUG-1998; 98US-0096768.
PR 17-AUG-1998; 98US-0096773.
PR 17-AUG-1998; 98US-0096791.
PR 17-AUG-1998; 98US-0096867.
PR 17-AUG-1998; 98US-0096891.
PR 17-AUG-1998; 98US-0096894.
PR 17-AUG-1998; 98US-0096895.
PR 17-AUG-1998; 98US-0096897.
PR 18-AUG-1998; 98US-0096949.
PR 18-AUG-1998; 98US-0096950.
PR 18-AUG-1998; 98US-0096959.
PR 18-AUG-1998; 98US-0096960.
PR 18-AUG-1998; 98US-0097022.
PR 19-AUG-1998; 98US-0097141.
PR 20-AUG-1998; 98US-0097218.
PR 24-AUG-1998; 98US-0097661.
PR 24-AUG-1998; 98US-0097951.
PR 26-AUG-1998; 98US-0097952.
PR 26-AUG-1998; 98US-0097954.
PR 26-AUG-1998; 98US-0097955.
PR 26-AUG-1998; 98US-0097971.
PR 26-AUG-1998; 98US-0097974.
PR 26-AUG-1998; 98US-0097978.
PR 26-AUG-1998; 98US-0097979.
PR 26-AUG-1998; 98US-0097986.
PR 26-AUG-1998; 98US-0098014.
PR 31-AUG-1998; 98US-0098525.
PR 16-SEP-1998; 98US-0100634.
PR 12-JAN-1999; 99US-0115565.
XX
XX (GETH) GENENTECH INC.
XX
XX Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;
PI Wood WI, Yuan J;
XX
XX WPI; 2000-072883/06.
DR P-PSDB; AAY66685.
XX
XX Membrane-bound proteins and related nucleotide sequences -
XX
XX Claim 2; Fig 136; 822pp; English.
XX
XX The invention provides membrane-bound PRO polypeptides and
CC polynucleotides encoding them. The PRO sequences of the invention were
CC identified based on extracellular domain homology screening. The PRO
CC sequences have homology with proteins including LDL receptors, TIE
CC ligands and various enzymes. The membrane-bound proteins and receptor
CC molecules are useful as pharmaceutical and diagnostic agents, Receptor
CC immunoadhesins, for instance, can be used as therapeutic agents to block

XX
PS
XX

Claim 2; Fig 161; 774pp; English.

Sequences AAS45925-AAS46231 represent DNA molecules encoding and PCR primers for PRO polypeptides of the invention. The sequences of the invention can be used to detect the presence of a tumour in a mammal by comparing the level of expression of a PRO polypeptide in a test sample of cells from the animal and a control sample of normal cells, whereby a higher level of expression in the test sample indicates the presence of a tumour in the animal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats and rabbits but are preferably human. The polypeptides can be used to stimulate tumour necrosis factor (TNF) alpha release from human blood, when contacted with it. A specific polypeptide can be used to stimulate the proliferation or differentiation of chondrocyte cells. The PRO proteins can be used to determine the presence of tumours and also susceptibility to tumour development, particularly adrenal, lung, colon, breast, prostate, rectal, cervical, or liver tumours, in mammalian subjects. The oligonucleotide probes specific for the PRO nucleic acids can be used for genetic analysis of individuals with genetic disorders.

XX
SQ

Sequence 2095 BP; 629 A; 382 C; 418 G; 666 T; 0 other;

Query Match 4.7%; Score 130.4; DB 22; Length 2095;
Best Local Similarity 48.9%; Pred. No. 1.4e-20;
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;

QY 511 TTACAGAAAGAGCGGAAGCTTCTTAAGCTCCAGATACAGACTGCAGGCGACACCTC 570
DB 494 TTACAGACAAGACTTTCACCTTCACACTTCGAGAGCATTCAAACTGCTCTCATCAAAATC 553
QY 571 CTTCTCTGCTCGCTGGTGTACCTCATCCCAAAACAGTTGGCTGAGCGGATGCCCATCC 630
DB 554 CATTTTGGTTCATTTGGTGACCTCCACCTTCAGATGTGAAGCCAGGCGAGCCCATTA 613
QY 631 GCGACAGCTGGGGAAGAGAGAGATGGTGAAGGAGAGAGCTCAAGACATCTTCCTCC 690
DB 614 GAGTTACTGGGTGAAAAAAGTCTTGGTGGGATGATGAGTCTTTACATTTTCTTAT 673
QY 691 TGGGGACACACAGCTGCGAGCGAAACGAAAGAGG-----TGGACGAGGAGAGCC 741
DB 674 TAGGCCAAGAGGCTGAAAGGAAGACAAATGTTGGCATGTTCTTAGAGATGAACACC 733
QY 742 AGCGACAGGGGACATATATCCAGAGGATTTCTAGACGTCTATTACAATCTGACCCCTGA 801
DB 734 TTTCTTTATGGTGACATAATCCGACAAAGATTTTATAGACACATATAATAACCTGACCTGA 793
QY 802 AGACCATGATGGCATAGATGGTCCATCGCTTTTGTCTCAGCGGCGCTTTGTATGA 861
DB 794 AAACCATATGGCATTCAGTGGGTAACTGAGTTTGGCCCAATGCGCAAGTACGTAATGA 853
QY 862 AACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACCTGCTTCTGAAGAAA 921
DB 854 AGACAGACACTGATGTTTTCATCAATGCTGCAATTTAGTGAAGTATCTTTTAAACCTAA 913
QY 922 ACAGAACACACAGGTTTTCACCTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGAGC 981
DB 914 ACCACTCAGAGAGTGTTCCTCAGAGGTTATCTCTTAATGATATTTATCTATAGAGAT 973
QY 982 CATTCAGCAAGTGTGTTGTGCTAGTAATCTGAATATCCGTTGGGACAGTACCCACCATCT 1041
DB 974 TTTACAAAACCCATATTTCTACAGAGATATCTTCAAGGTGTTCCCTCCATATC 1033
QY 1042 GTCGCGGACCGGTACGTGTTTCTTGGCGAGCTGCGAGTCAAGTGTACAATGTCTCCA 1101
DB 1034 GCAGTGGGTGGTGTATATAATGTCCAGAGATTTGTGTCGAAGATCTATGAATGATGG 1093
QY 1102 AGAGGCTCCATACATTAACCTGGAAGAGCTGTTTGTGGGGCTCTCCCTCGAAAGCTGA 1161
DB 1094 GTCAGTAAACCCCATCAAGTTTGAAGATGTTTATGTGCGGATCTGTTTGAATTTATTA 1153
QY 1162 ACATCAGATTGGAGAGCTCCACTCCAGCCGACCTTTTTTCCAGGGGCTTACGTTCT 1221
DB 1154 AAGTGAACATTCATATCCAGAGACACAAATCTTTTCTTATATAGAATCCATTTGG 1213

QY 1222 CCGTATGCTCTCTCAGGAGGATCGTGCGCTGCCACTTCATCAAGCCTCGGACTCTCTTGG 1281
DB 1214 ATGCTGTCAACTGAGACGTGTGATTGCGCCCATGGCTTTTCTTCCAAGGAGATCATCA 1273
QY 1282 ACTACTGGCAGG 1293
DB 1274 CTTTGTGGCAGG 1285

RESULT 10

AAF92075

ID AAF92075 standard; cDNA; 2095 BP.

XX

AC AAF92075;

XX

DT 15-MAY-2001 (first entry)

XX

DE Human PRO1074 cDNA.

XX

KW Human; PRO protein; mapping; ss.

XX

OS Homo sapiens.

XX

PN WO200116318-A2.

XX

PD 08-MAR-2001.

XX

PF 24-AUG-2000; 2000WO-US23328.

XX

PR 01-SEP-1999; 99WO-US20111.

PR

15-SEP-1999; 99WO-US21090.

PR

07-DEC-1999; 99US-0169495.

PR

09-DEC-1999; 99US-0170262.

PR

11-JAN-2000; 2000US-0175481.

PR

18-FEB-2000; 2000WO-US04341.

PR

18-FEB-2000; 2000WO-US04342.

PR

22-FEB-2000; 2000WO-US04414.

PR

01-MAR-2000; 2000WO-US05601.

PR

03-MAR-2000; 2000US-0187202.

PR

23-APR-2000; 2000US-0199397.

PR

25-MAY-2000; 2000WO-US14042.

PR

05-JUN-2000; 2000US-0209832.

XX

(GETH) GENENTECH INC.

XX

Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;

PI

Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;

XX

WPI: 2001-183260/18.

DR

P-PSDB; AAB87543.

XX

Eighty four nucleic acids encoding PRO polypeptides, useful in

PT

molecular biology, including use as hybridization probes, and in

XX

chromosome and gene mapping.

XX

Claim 2; Fig 35; 278pp; English.

PS

XX

The present sequence is the coding sequence for a human PRO polypeptide (secreted and transmembrane). The PRO protein, and PRO agonists, PRO antagonists or anti-PRO antibodies are useful for preparation of a medicament useful in the treatment of a condition which is responsive to the PRO protein, agonists, antagonists or anti-PRO antibodies. The PRO protein may also be employed as molecular weight markers for protein electrophoresis. The PRO coding sequence has applications in molecular biology, including use as hybridisation probes, and in chromosome and gene mapping.

SQ Sequence 2095 BP; 629 A; 382 C; 418 G; 666 T; 0 other;

Query Match

Best Local Similarity 48.9%; Score 130.4; DB 22; Length 2095;

Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;

CC and treatment of disorders including infections by bacteria, yeasts and
CC viruses (including HIV) and protozoa; metabolic and hormonal disorders;
CC immune disorders (including severe combined immunodeficiency (SCID)
CC and AIDS; thrombosis; cancer; and traumatic or surgical wounds.

Query Match 4.58; Score 124; DB 20; Length 2420;

Query Match
Best Local Similarity
49.08; Pred. No. 4.7e-19;
Score 124, DB 20.

| Seq. ID | Accession | Length | Score | E-value | Identical | Similarity | Positives | Conservative | Mismatches | Indels | Gaps |
|---------|-----------|--------|-------|---------|-----------|------------|-----------|--------------|------------|--------|------|
| 1 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 2 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 3 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 4 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 5 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 6 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 7 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 8 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 9 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 10 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 11 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 12 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 13 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 14 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 15 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 16 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 17 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 18 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 19 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 20 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 21 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 22 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 23 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 24 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 25 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 26 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 27 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 28 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 29 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 30 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 31 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 32 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 33 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 34 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | 100 | 0 | 0 | 0 |
| 35 | U00096.1 | 430 | 100.0 | 0.0 | 100 | 100 | 100 | | | | |

| | | | |
|----|------|--|------|
| Qy | 551 | AGACTGCAGCGACACACCTTCCCTTCCTGCTGCTGTGGTACCCTCATCCCACAACACAGTT | 610 |
| Dy | 1129 | AAATGCCAAGAAGAAAAGTCCCTTTTTTAATAGTACTATACTAGCTGCAGAGCCTGCACAAAT | 1188 |
| Qy | 611 | GGCTGAGCGCATGGCCATCCGGCAGACAGCTGGGGGAAAAAGAGAGATTGGTGAAGGGAAAAAGCA | 670 |
| Dy | 1189 | AGAGCTAGAAGAGCTATTTCGGCAAACTTGGGGCAATGNAAGCTTAGCACCTGGTATTCA | 1248 |
| Qy | 671 | GCTGAAGACATTCCTCCCTCTGGGGACCAACA - GCAGTGCGAGCGGAAACGAAGAGGTGG | 729 |
| Dy | 1249 | AATCACAAGAAATATTTTGTGTGGGCTTAAGTATTAAAGCTAAATGGCTACCTTCAACGTCG | 1308 |
| Qy | 730 | A-----CCAGSGAGCAGCAGCAGCGGGGACACATTATCCAGAGGAGTTTCCTTAGACGCTCA | 784 |
| Dy | 1309 | AATACTGGAAGAAGCAGACATATCATGRTATTAATCAACAGGAATACTTAGATACGTA | 1368 |
| Qy | 785 | TTACAATCTGACCCTGAAGACCATGATGGGCATAGAATGGGTCCCATCGCTTTTGTCTCTCA | 844 |
| Dy | 1369 | CTATAATTTGACCATTAAAACACTAATGGCATGAATGGGTGTCACATACTGTCACCA | 1428 |
| Qy | 845 | GGCGGCGTTTGTCATGAAAACAGACTCAGACATGTTTCATCAATGTTTGACTATCTGACTGA | 904 |
| Dy | 1429 | TATTCACATATGTTATGAAAACACTGACAGTGACATGTTGTCAACACTGAAATATTTAAATCAA | 1488 |
| Qy | 905 | ACTGCTTCTCAAGAAAAAACAGAACCAACAG ----- GTTTTCTACTGCGCTTCTTGAACCT | 958 |
| Dy | 1489 | TAAGTTACTGAAGCCAGATCGCTCCACAGACATAACTATTCTACTGGTTACCTAATGCG | 1548 |
| Qy | 959 | CAATGAGTTTCCCATCAGCGAGGCATTACGAAGTGGTTTGTGAGTAAATCTGAATATCC | 1018 |
| Dy | 1549 | AGGATATGSCACCCCAATCGAAACAAAGATAGCAAGTGTACATGCCACCAGACCTCTACCC | 1608 |
| Qy | 1019 | GTGGGACAGGTACCCACCATTCTGCTFCGGCACCGCTACGTGTTTCTCGCGACGTGGC | 1078 |
| Dy | 1609 | AAGTGAGCGTTATCCTGCTCTTGTTCTGGAACCTGTTATGTTTCTTGAGATCTGGC | 1668 |
| Qy | 1079 | GAGTCAGGTTACAATGTCTCCAAGAGCCTCCCATACATTAAACTTGAAGACGTGTTGT | 1138 |
| Dy | 1669 | AGAAAAGATTTTTAAAGTTCTTTTAGTTATCCGCCGCTTGGCATTTGGAAGATGATATGT | 1728 |
| Qy | 1139 | GGGCTCTGCCTCGAAGGCTGNACATCAGATTGGAGAGCTCCACTCCCACGCGACCTT | 1198 |
| Dy | 1729 | AGGGATCTGTCTTGGCCAAAGTTGAGAAATGATCCTGTACCCCTCCCAATGAGTTGTGTT | 1788 |
| Qy | 1199 | TTTTCCAGGGGGCTTACGGTTCTCCGATGCGCTTTCAGGAGGATCGTGGCCTGCCACTT | 1258 |
| Dy | 1789 | CAATCACTGCGAGTCTCTATTTCGAGCTGTAATAACAGCCACCTAATTAACCTCTCATCA | 1848 |
| Qy | 1259 | CATCAAGCCTCGGACTCTCTTGACTACTGTCAGCGCTCTAGAGAAATCCCCGGGGGAAGA | 1318 |
| Dy | 1849 | GTTCACGCCCTAGTGAACCTGATAAAATPACTTGAACCACTTTACAACAAAAATAAGCACATGC | 1908 |
| Qy | 1319 | TTGTCGCCCTGTCTGCTGGGAGCCACAGAGGCACATCCCGACAAGTTTCAGATAACCCGCTG | 1378 |
| Dy | 1909 | CTGTGCCACGCAGCAAAAGAAAGCAGCGAGGTATCGCCACCCTAAACTACATTTAGAA | 1968 |
| Qy | 1379 | GGGATGTTTTTCTAGATTTTGGAA | 1404 |
| Dy | 1969 | AAGACAAATTTTTTTTCAAATGTGCAA | 1994 |

Search completed: April 11, 2003, 20:07:01

Job time : 520.132 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 15:35:54 ; Search time 3089.38 Seconds
(without alignments)
14547.424 Million cell updates/sec

Title: US-09-914-152-2
Perfect score: 2775
Sequence: 1 gtgaattctcttctctgc.....aaaaaaaaaaaaaaaaaaaaa 2775

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pln:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_mam:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| C 1 | 505.4 | 18.2 | 507 | 9 | AI739518 |
| C 2 | 455 | 16.4 | 455 | 10 | AW136254 |
| C 3 | 374.4 | 13.5 | 434 | 12 | BF748181 |
| C 4 | 326 | 11.7 | 338 | 9 | AJ003597 |
| C 5 | 238.4 | 8.6 | 300 | 9 | AU100432 |
| C 6 | 234 | 8.4 | 400 | 17 | AZ768940 |

| | | | | | |
|------|-------|-----|------|----|-----------|
| 7 | 234 | 8.4 | 528 | 17 | AQ266346 |
| C 8 | 184 | 6.6 | 1101 | 17 | CNS05BT1 |
| 9 | 179 | 6.5 | 636 | 13 | BG965086 |
| 10 | 135 | 4.9 | 1514 | 11 | AK003837 |
| C 11 | 133.6 | 4.8 | 491 | 14 | BQ021377 |
| 12 | 133.6 | 4.8 | 3199 | 11 | BC013258 |
| 13 | 132.2 | 4.8 | 596 | 13 | BJ501651 |
| 14 | 128.8 | 4.6 | 3564 | 11 | BC028571 |
| 15 | 127 | 4.6 | 233 | 10 | BB595953 |
| 16 | 123.4 | 4.4 | 444 | 13 | BM403751 |
| C 17 | 122.8 | 4.4 | 400 | 9 | AI1311905 |
| C 18 | 122.8 | 4.4 | 462 | 9 | AI742535 |
| C 19 | 122.8 | 4.4 | 525 | 9 | AI820087 |
| 20 | 122.6 | 4.4 | 879 | 14 | BQ895238 |
| C 21 | 122.4 | 4.4 | 329 | 9 | AI183500 |
| C 22 | 121 | 4.4 | 651 | 14 | BQ180890 |
| 23 | 120.8 | 4.4 | 399 | 17 | AQ475568 |
| 24 | 120.4 | 4.3 | 701 | 12 | BF131199 |
| 25 | 120 | 4.3 | 688 | 9 | AF063529 |
| 26 | 120 | 4.3 | 688 | 9 | AI110655 |
| 27 | 117.2 | 4.2 | 965 | 17 | CNS05182 |
| C 28 | 116.8 | 4.2 | 558 | 14 | BQ363849 |
| C 29 | 116.2 | 4.2 | 690 | 17 | AG130187 |
| 30 | 116 | 4.2 | 343 | 12 | BF759672 |
| C 31 | 115.4 | 4.2 | 1021 | 17 | CNS04474 |
| C 32 | 114.8 | 4.1 | 574 | 10 | BE652985 |
| C 33 | 113.4 | 4.1 | 664 | 17 | AQ346244 |
| 34 | 113.2 | 4.1 | 220 | 14 | F05486 |
| 35 | 112.6 | 4.1 | 475 | 17 | AQ195633 |
| 36 | 112.6 | 4.1 | 491 | 17 | AQ282423 |
| 37 | 112.4 | 4.1 | 247 | 14 | T08840 |
| C 38 | 112.4 | 4.1 | 632 | 17 | AQ277749 |
| 39 | 112 | 4.0 | 664 | 14 | BQ574062 |
| 40 | 112 | 4.0 | 725 | 14 | BQ448351 |
| 41 | 111.6 | 4.0 | 730 | 17 | AG029969 |
| 42 | 110.2 | 4.0 | 380 | 17 | AQ418581 |
| 43 | 110.2 | 4.0 | 417 | 17 | AQ415767 |
| 44 | 110.2 | 4.0 | 586 | 17 | AQ372907 |
| C 45 | 109.6 | 3.9 | 574 | 10 | BE272364 |

ALIGNMENTS

RESULT 1
AI739518/c
LOCUS AI739518 507 bp mRNA linear EST 20-DEC-1999
DEFINITION w123a06.x1 NCI_CGAP_Col6 Homo sapiens cDNA clone IMAGE:2391058 3', mRNA sequence.
ACCESSION AI739518
VERSION AI739518.1 GI:5101499
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 507)
AUTHORS NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
Cloned distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1331 Std Error: 0.00
Seq primer: -40UP from Gibco

AI739518
w123a06.x1 NCI_CGAP_Col6 Homo sapiens cDNA clone IMAGE:2391058 3', mRNA sequence.
AI739518
AI739518.1 GI:5101499
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
Cloned distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1331 Std Error: 0.00
Seq primer: -40UP from Gibco

Db 395 CTCAGAGCCCTAAAGTGGGCGCTGGTGAAGCAGGGTGGTCTCGGTCCACCTTCCCAAGCC 336
Qy 2413 TGAGCCAGGCTCATCTTCATTGAATGCTCATTTGGCGGAGGAACAACACTGAACCTTTGGG 2472
Db 335 TGAGCCAGGCTCATCTTCATTGAATGCTCATTTGGCGGAGGAACAACACTGAACCTTTGGG 276
Qy 2473 TTTGCTGTTAGCCTTCAGTTTGTCCGCTGCCCTCCCTACCCAGAGGTTTGTGCGAGCCTG 2532
Db 275 TTTGCTGTTAGCCTTCAGTTTGTCCGCTGCCCTCCCTACCCAGAGGTTTGTGCGAGCCTG 216
Qy 2533 TGTTCAGAGGTTGTATAAACCAAGGTACTTCGTTAGTTTGGCCCATTCACCCATGGTCA 2592
Db 215 TGTTCAGAGGTTGTATAAACCAAGGTACTTCGTTAGTTTGGCCCATTCACCCATGGTCA 156
Qy 2593 CGTCACATGCAAACTAATCTTCGCTCCCTAATATAGAAATGATTTTCTTTTAAATTTTAA 2652
Db 155 CGTCACATGCAAACTAATCTTCGCTCCCTAATATAGAAATGATTTTCTTTTAAATTTTAA 96
Qy 2653 CTTTACCAGACTTTTACCTTTGTACAGAGAGAGGCTTCACATGCTGTGTACATATAA 2712
Db 95 CTTTACCAGACTTTTACCTTTGTACAGAGAGAGGCTTCACATGCTGTGTACATATAA 36
Qy 2713 ATGTTGACTAACTCTTAAAAAATAAAAAA 2747
Db 35 ATGTTGACTAACTCTTAAAAAATAAAAAA 1

RESULT 3
BF748181/c
LOCUS RC3-BN0411-021000-021-c04 BN0411 Homo sapiens cdNA, linear EST 10-JAN-2001
DEFINITION BF748181
ACCESSION BF748181
VERSION BF748181.1 GI:12074857
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 434)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=RC3&t2=RC3-BN0411-021000-021-c04&t3=2000-10-02&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 24
High quality sequence stop: 432.
Location/Qualifiers
1. 434
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BN0411"
/dev_stage="Adult"
/note="Organ: breast normal; Vector: puc18; Site:1: Sma1; Site:2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application

BASE COUNT 104 a 117 c 130 g 83 t
ORIGIN
Query Match 13.5%; Score 374.4; DB 12; Length 434;
Best Local Similarity 97.2%; Pred. No. 9e-55; Indels 0; Gaps 0;
Matches 381; Conservative 0; Mismatches 11;
Qy 959 CAATGAGTTTCCCATCAGGAGCCCATTCAGCAAGTGGTTGTGTCAGTAAATCTGAATATCC 1018
Db 434 CAATGAGTTTCCCATCAGGAGCCCATTCAGCAAGTGGTTGTGTCAGTAAATCTGAATATCC 375
Qy 1019 GTGGACAGGTACCCACCATTCGTCGCGCACCGGCTACGTGTTTTCTGGGACGTGGC 1078
Db 374 GTGGACAGGTACCCACCATTCGTCGCGCACCGGCTACGTGTTTTCTGGGACGTGGC 315
Qy 1079 GAGTCAGGTGTACAATGTCTCCAAGAGGCTCCCATACATTAACCTGGAAGAGCTGTTGT 1138
Db 314 GAGTCAGGTGTACAATGTCTCCAAGAGGCTCCCATACATTAACCTGGAAGAGCTGTTGT 255
Qy 1139 GGGCTCTCGCTCGAAAGGCTGAACATCAGATTGGAGGAGCTCCACTCCAGCCGACCTT 1198
Db 254 GGGCTCTCGCTCGAAAGGCTGAACATCAGATTGGAGGAGCTCCACTCCAGCCGACCTT 195
Qy 1199 TTTTCCAGGGGCTTACGCTTCCGTTATGCTCTTTCAGAGGAGTGGTGGCTGGCCTT 1258
Db 194 TTTTCCAGGGGCTTACGCTTCCGTTATGCTCTTTCAGAGGAGTGGTGGCTGGCCTT 135
Qy 1259 CATCAAGCTCGGACTCTTTCGACTACTTGGAGGCTCTAGAGATTCCCGGGGGAAGA 1318
Db 134 CATCAAGCTCGGACTCTTTCGACTACTTGGAGGCTCTAGAGATTCCCGGGGGAAGA 75
Qy 1319 TTGTCGCGCTGTCTGAGGGGAGCCAGGCA 1350
Db 74 TTGTCGCGCTGTCTGAGGGAGCCCGGAGGCA 43

RESULT 4
AJ003597/c
LOCUS AJ003597 Selected chromosome 21 cdNA library Homo sapiens cdNA
DEFINITION AJ003597 338 bp mRNA linear EST 04-DEC-1997
ACCESSION AJ003597
VERSION AJ003597.1 GI:2578270
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 338)
Szulzewsky,I., Hunt,E., Nguyen,M., Korn,B., Roehrdanz,B., Lehrach ,H. and Yaspo,M.L.H.
TITLE An integrated transcript map for the whole human chromosome 21
JOURNAL Unpublished (1997)
COMMENT Contact: Yaspo, M.-L.
Max Planck Institut fuer Molekulare Genetik
Innestrasse 73, D14195 Berlin-Dahlem, Germany.
Location/Qualifiers
1. 338
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="21q"
/clone="MPI19-12J9"
/clone_lib="Selected chromosome 21 cdNA library"
/note="Swaroop et al. (1991) Nucleic Acids Res. 19: 1954."
BASE COUNT 76 a 92 c 82 g 87 t 1 others
ORIGIN
Query Match 11.7%; Score 326; DB 9; Length 338;
Best Local Similarity 98.2%; Pred. No. 2.1e-46;

| | | | |
|------------|--|--|----------------------------|
| Db | 40 | GTGAATTCCTCTTCTCTGCTGGAGCTGGGATATACTTCTNCTCCTCGCCCTTGGACATCA | 99 |
| Qy | 61 | GAGCTGCAGGCTCTCTGGCTTTGGACCCGAGGATTTATACCAAGCAGGTTTCTTGGGTTCT | 120 |
| Db | 100 | TAGCTGCTGGCTCTCTGGNCTTTGGACCCGAGGATTTATACCAAGCAGGTTTCTTGGGTTCT | 159 |
| Qy | 121 | TCAGGCCCTTTGGCCTTGGACTGATAGTTACACCATTTGGCATATCTGTTCTTGAGGCTCTT | 180 |
| Db | 160 | TCAGGCCCTTTGGCCTTGGACTGATAGTTACACCATTTGGCATATCTGTTCTTGATGCTT | 219 |
| Qy | 181 | GGTCTTGACTTGAGCCACACTCTTGGCATCCAGCGTCTCCAGCTTGCATGGCTGTGCAC | 240 |
| Db | 220 | GGTCTTGANNATTCACACTTNGCATCCACGCTCCAGCTTGCATGGCTGTGCAC | 279 |
| Qy | 241 | GTGATTCCTGTCAGAAATCACC | 261 |
| Db | 280 | GTGATTCCTGTCAGAAATCACC | 300 |
| RESULT 6 | | | |
| LOCUS | AZ768940/c | | |
| DEFINITION | 1M0569K09F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0569K09 F, DNA sequence. | | |
| ACCESSION | AZ768940 | 400 bp | DNA linear GSS 16-FEB-2001 |
| VERSION | AZ768940 | | |
| KEYWORDS | AZ768940.1 | GI:12888559 | |
| SOURCE | GSS. | | |
| ORGANISM | Mus musculus house mouse. | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus; 1 (bases 1 to 400) | | |
| AUTHORS | Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Stokes,R., Tingey,A., von Niederhausern,A., and Wright,D.,Weiss,R. | | |
| TITLE | Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts | | |
| JOURNAL | Unpublished (2000) | | |
| COMMENT | Contact: Robert B. Weiss University of Utah Genome Center University of Utah Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLCT, UT 84112, USA Tel: 801 585 5606 Fax: 801 585 7177 Email: ddunn@genetics.utah.edu Insert Length: 10000 Std Error: 0.00 Plate: 0569 row: K column: 09 Seq primer: CGTTGTAAACAGCGCCAGT Class: plasmid ends High quality sequence stop: 400. Location/Qualifiers 1. .400 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="UUGC1M0569K09" /clone_lib="Mouse 10kb plasmid UUGC1M library" /sex="Male" /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-" /note="Vector: pWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnates/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. The | | |
| FEATURES | source | | |

JOURNAL Submitted (12-APR-2000)
COMMENT This sequence is a single read and was generated as part of a large scale clone-and-sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/tetraodon>.

FEATURES
source Location/Qualifiers
1. .1101
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="013p23"
/clone.lib="C"
/note="Genoscope sequence ID : COCC013CH12T1-end : T3"

BASE COUNT 235 a 239 c 325 g 283 t 19 others

ORIGIN
Query Match 6.6%; Score 184; DB 17; Length 1101;
Best Local Similarity 58.5%; Pred. No. 3.4e-22;
Matches 354; Conservative 0; Mismatches 242; Indels 9; Gaps 2;
Qy 568 CTCCTTCTCTGCTGCTGACCTCATCCCAACACAGCTGGCTGAGCCATGGCCA 627
Db 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
663 CTCCTTCTGCTCATCTCATCAGCACACGACGAGGTTTGACGCGCGCAGGCCA 604
Qy 628 TCCGCGACAGTGGGGGAAAGAGAGGATGGTGAAGGAAAGCAGCTGAAGACATCTTCC 687
Db 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
603 TCCGGAGACCTGGGGGACGAGACACCTTCACGAGCTCGCATCTTCACCGCTTTC 544
Qy 688 TCTCTGGGACCAACAGCAGTGTGAGGAAAGAGAGTGTGACGAGGAGCCAGCCGAC 747
Db 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
543 TNCCTGGGAGGAACACCATGAAGTCTCAACCCAGATGCTGACGAGGAGGACGATCT 484
Qy 748 ACGGGGACATTCACAGAGGATTTCTTAGAGCTCTATTACAACTTGACCTTGAGACCA 807
Db 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
483 TCCATGACATCGTATGAGAACTTCATCGATTCCTACCAACACCTCAGCTCAAACTT 424
Qy 808 TGATGGGATAGAAATGGGTCCATCGCTTTTGTCTCAGCGCGCTTTGTGATGAAAAACAG 867
Db 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
423 TGATGGGATAGCGTGGGTGGCCACTTCTGCCCAAGCTCAATATGTGATGAAGACAG 364
Qy 868 ACTAGACATGTTCAATATGTTGACTATCTGACTGAAGTGTCTCTGAAG-----AAAA 921
Db 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
363 ACAGCGACATCTTCGCTCAACATGNAACATCTNATCAAGCTCTCTGAAGCCGACGACCA 304
Qy 922 ACAGAACACACAGCTTTTCACTGGCTTCTGAACTCAATAGTTCCTCATCAGGCAGC 981
Db 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
303 ACCCGAGGAGGAGTATTCACGGCTACGTTA---TAAACGGCGACCGATACGGGACA 247
Qy 982 CATTCAGCAAGTGTGTCAGTAATCTGAATATCCCTGGGACAGGTACCCACCATTTCT 1041
Db 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
246 TCGCGAGTAAAGTGTACATGTGAGGGACTTGTATCCCGAGAGCAATACCCGCGTTCT 187
Qy 1042 GCTCGGACCGGCTACGTGTTTCTGGGAGCTGGCGAGTCAAGTGTACAATGTCTCCA 1101
Db 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
186 GCTCANGCACCNGTACGTGTTCTCAGCGGAGCTGGCGGAGTCTATATAAGACTCC 127
Qy 1102 AGAGGCTCCCATACATTAACCTGGAAGAGCTGTTGTGGGGCTCTGCCTCGAAAGCTCA 1161
Db 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
126 TGCACACAGACTGTGCACCTGGAGGAGCTGTACNTGGACTGTGTTGCATAGCTGG 67
Qy 1162 ACATC 1166
Db 1111
66 GCATC 62

RESULT 9
BG965086
LOCUS 602829157F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4983830 5',
DEFINITION mRNA sequence.
ACCESSION BG965086
VERSION BG965086.1 GI:14352723
KEYWORDS EST.
SOURCE house mouse.

ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 636)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM10989 row: a column: 15
High quality sequence start: 17
High quality sequence stop: 634.

FEATURES
source Location/Qualifiers
1. .636
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4983830"
/clone.lib="NCI_CGAP_Co24"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 155 a 172 c 155 g 154 t

ORIGIN
Query Match 6.5%; Score 179; DB 13; Length 636;
Best Local Similarity 66.2%; Pred. No. 3.2e-21;
Matches 321; Conservative 0; Mismatches 155; Indels 9; Gaps 4;
Qy 335 ATTTCCTCTTGCACTTTACACTGTGGCTTTAGCTTTTCAAAACAGAGGTTCTCTTACCCA 384
Db 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
159 ATTTCCTCTTGCTTCTATCTCTGTGGTTTCATCTCTCGGACGAACGTGAGCTCTTACTCG 218
Qy 395 GCAAAAATGGCTTCCCGAAGATGAGATGATGATATTTGCCCTTCTGGTCTGGGGC 454
Db 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
219 GC-AAAGTGGCTCACATGAAGACAAGGCTAGTTTACGGCTCCATTTCTCATGATGGCGC 277
Qy 455 TCTTTGTTTGTATTTAGCATGTACAGTCTAAATCTTTTCAAGAACAGCTCTTTGTTTA 514
Db 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
278 ACTCTGTTGTACTTCAGCATGGATCTTTCA-----GAGAACTCCCGTTGTTTAAAC 331
Qy 515 CAAGAAAGACGGGAACCTTCTTAAGCTCCAGATA-CAGACTGCAGGAGACACCTCCCT 573
Db 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
332 GAAAAGTCAACGGAGTCTCTTCAGATTCGGATATAAGACTGCAAGCAGAACGCCCTT 391
Qy 574 TCTCTGCTCTGCTGGTGAACCTCATCCCAACAGAGTTGGCTGAGCGCTGGCCATCCGC 633
Db 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
392 TCTGCTGCTGCTGGTGAACCTCATCTCAGAGCAGCTGGCCGCTGCATGGCCATCCGCA 451
Qy 634 AGACGTGGGGGAAAGAGAGGATGGTGAAGGG-AAAGCAGCTGAAGACATTTCTTCCTCTG 692
Db 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
452 AGACGTGGGGTAGAGAGACATCTGTGAGGGCACAACAGGTGAGGACCTTCTTCCTCTCTG 511
Qy 693 GGGACCAACAGCTGCAGCGGAACGAAGAGGTGGACCAAGAGGAGCGGACGACGGG 752
Db 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
512 GGGACCTCCGACAGCAGCCAGGAGATGGACGCCAACAAACCCCTGGAGAGGAGCAGCAGCG 571
Qy 753 GACATTATCCGAAGGATTTCTTAGACGCTATTACAATCTGACCTTGAGACCATGATG 812
Db 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
572 GACATTATCCGAAGGACTTCAAGGATGCTTAACCTTCAACTGACCTGAAGACCATGATG 631
Qy 813 GGAT 817
Db 1111
632 GGTAT 636

65 AGAGTTACTTGGGTGAGAAAAGTCCTGGTGGGATATGAGTGCTACGCTTTTCTTCA 124
QY 690 CTGGG--GACCACAGCAGTCAGCGGAACGAAGAGGTGACGAGGAGCCAGCGCA 746
Db 125 CTAGGCCAGCGCTGAAGGAAGACAAACGTTAGCCCTCTCTCTGGAGGATGAGCAC 184
QY 747 C-----ACGGGACATATCCAGAGGATTTCTTAGAGCTATTATACATCTGACCCCTG 800
Db 185 GTTCTCTATGGTGATATTACGGCAAGACTTTCTAGACACATATAATAAATTAACCTTG 244
QY 801 AAGACCATGAGGATAGATAGTGGTCCATCGCTTTTGTCCCTCAGCGCGCTTTGTGATG 860
Db 245 AAAACCATATGCGCTTTCAGGTGGGTAAATGAGTATTTGCCCAATGCCAAGTATATCATG 304
QY 861 AAAACGACTCAGACATGTTCAATGTTGACTATCTGACTGAAGTCTCTGAGAA 920
Db 305 AAAACGACATGATGTTTTCATCAACACTGCAATTTAGTCAAGTAATCTTTAAACCTA 364
QY 921 AACAGAACACAGGTTTTTCACTGGCTTTTGAACCAATGAGTTTCCCATCAGGCGAG 980
Db 365 AACCACTCAGAGAAGTTTTTCAAGGCTATCTCTAATTTGATAAATTTCTCTATAGAGA 424
QY 981 CCATTGACAGCAAGTGTGTTGTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTC 1040
Db 425 TTTTCCATAAAACCATTCATACAGAGTACCCCTTCAAGGTGTTCCCTCCCTAC 484
QY 1041 TGCTCCGACCGGCTAGCTGTTTTCTGGGACGCTGGGAGTGACAGTGTACAAATCTCTCC 1100
Db 485 TGCAGCGGCTGGCTACATATATGTCGGGACCTGTCGCCAGGCTCTACGAGATGATG 544
QY 1101 AAGACGCTCCCATACATTAACATGGAAGACGTTGTTGGGGCTCTGCTCCAAAGGCTG 1160
Db 545 AGTCAGCTGAAGCCCATCAAGTTTGAAGACGTTTATGTTGGCATCTGTTGAAATTTGTA 604
QY 1161 AACATCAGATTGGAGGAGCTCCACCTCCAGCGAGCTTTTTCAGGGGCTTACGCTTC 1220
Db 605 AAAGTGGACATTCATATCCAGAACACAAACCTTTTCTTCTGTACAGATCCACTTG 564
QY 1221 TCCGTATGCTCTCAGGAGGATCGTGGCTGCCACTTCATCAAGCCTCGGACTCTCTTG 1280
Db 665 GATGATATGACTCAGACGCGTGATGCGCCCATGCTTTCTTCCAAAGGAGATCATC 724
QY 1281 GACTACTGCGAGG 1293
Db 725 ACATTCTGCGAGG 737

RESULT 11
BO021377/c
LOCUS BO021377 491 bp mRNA linear EST 27-MAR-2002
DEFINITION UI-H-DHI-axf-p-12-0-UI.s1 NCI-CGAP_DHI Homo sapiens cDNA clone
IMAGE:5828219 3', mRNA sequence.
ACCESSION BO021377
VERSION BO021377.1 GI:19756655
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 491)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
UNPUBLISHED (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Cloning Distribution: Clome Distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
The following repetitive elements were found in this cDNA

sequence: 32-98, >MLT2D#LTR/Retroviral (matched complement) 32-76,
>MLT2B2 (matched complement) 137-323, >MLT2B2 (matched complement)
250-465, >MLT2B2 (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source

Location/Qualifiers

1. 491
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5828219"
/clone_lib="NCI-CGAP_DHI"
/tissue_type="Metastatic Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
NCI-CGAP_DHI is a normalized cDNA library containing the
following tissue(s): VS-8 Cell line from Metastatic
Chondrosarcoma in Lung. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is AGATCATTCG.

TAG_LIB=UI-H-DHI

TAG_TISSUE=Lung

TAG_SEQ=AGATCATTCG

BASE COUNT 144 a 93 c 123 g 131 t

Query Match 4.8%; Score 133.6; DB 14; Length 491;
Best Local Similarity 78.9%; Pred. No. 2.3e-13;
Matches 172; Conservative 0; Mismatches 44; Indels 2; Gaps 1;

QY 12 TTTCTCTGCTGGAGCTGGGATATCTCTCTCTCTCCCTTGGACATCAGAGCTCGAGCG 71

Db 420 TCTCTCTCTGGAGCTGGGATATCTCTCTCTCTCCCTTGGACATCAGAGCTCGAGCG 361

QY 72 TCTCTGGCCTTTGGACCGGAGGATTTATACCAAGAGGTTTCTGGGTTCTCAGGCTTTG 131

Db 360 TCTCCGGCCTTTGAACTTCAGGACTTGACCAAG--AGGCCCTGGGTTCTCAGGCTTTG 303

QY 132 GCCTTGGAGCTAGATTACACCATTCGCATATCTGCTCTGAGGCTCTTGGTCTTGACT 191

Db 302 GCTTTGGAGCTAGAGTTACACAATCAGCTTCCCTGGTCTTGGAGCTTTTTCAGACTTAACT 243

QY 192 GAGCCACACTCTCGGATCCGAGCGCTCTCCAGCTTGCA 229

Db 242 GAGCCATGCTACCAGCATCCAGGGTCTCCAGCCTACA 205

RESULT 12

BO013258

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BO013258 Homo sapiens, clone IMAGE:3867325, mRNA.

BO013258 Homo sapiens, clone IMAGE:3867325, mRNA.

BO013258.1 GI:15301554

HTC.

Homo sapiens.

Homo sapiens.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 3199)

Direct Submission

Strausberg, R.

Submitted (27-AUG-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

HTC 27-AUG-2001

FEATURES

source

further details.

Location/Qualifiers

1..233

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="C230001K09"

/clone_lib="RIKEN full-length enriched, 0 day neonate cerebellum"

/tissue_type="cerebellum"

/dev_stage="0 day neonate"

/lab_host="DH10B"

/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTCTTTTTTTTTTTN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot - 20.0 and subtraction to Rot - 479.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATCTCGAGTTAAATTAATCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT

57 a

57 c

74 g

44 t

1 others

ORIGIN

Query Match

Best Local Similarity 4.6%; Score 127; DB 10; Length 233;

Matches 166; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Qy 570 CCCTTCCTCCTGCTGGTGACCTCATCCACAAACAGTTGGCTGAGCGCATGGCCATC 629

Db 2 CCTTTCTGCTGGTGCTGACGCTCATCTCACAAGCAGCTGGCGCTCGCATGGGCATC 61

Qy 630 CGGCAGACGTGGGGAAGAGAGATGGTGAAGGGAAGCAGCTGAAGACATTTCTCTC 689

Db 62 CGGAAGAAGTGGGGTAAAGAGACATCTGTGGAGGGCCACACAGGTGAGGACCTCTCTT 121

Qy 690 CTGGGGACCAACGACGTGCGGGAAGAGAGAGTGGAGAGAGAGTGGAGAGAGAGAGAGAG 749

Db 122 CTGGGATCTCCGACAGGACCGAGGAATGGACCCACACACCCGTGGAAAGCGAGNAGTAC 181

Qy 750 GGGGACATATCCAGAGAGATTTCTTAGACGTCTATTACAACTGACCCCTGA 801

Db 182 GGGGATATTCCAGAGAGATTTCAAGGATGCCTTACTCCAAACGTGACCCCTGA 233

Search completed: April 11, 2003, 22:37:14

Job time : 3113.38 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 15:37:18 ; Search time 76.6272 Seconds
(without alignments)
11106.080 Million cell updates/sec

Title: US-09-914-152-2
Perfect score: 2775
Sequence: 1 gtgaattctcttctctgc.....aaaaaaaaaaaaaaaaaaaa 2775

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 133.6 | 4.8 | 3502 | 2 | US-08-724-394A-16 |
| 2 | 124.8 | 4.5 | 333 | 4 | US-08-905-223-218 |
| 3 | 114.8 | 4.1 | 1134 | 4 | US-09-482-180A-3 |
| 4 | 78.4 | 2.8 | 1420 | 4 | US-09-482-180A-1 |
| 5 | 68.6 | 2.5 | 1191 | 4 | US-09-459-133-3 |
| 6 | 67 | 2.4 | 1167 | 4 | US-09-459-133-14 |
| 7 | 61 | 2.2 | 2550 | 6 | 5258287-23 |
| 8 | 58.2 | 2.1 | 703 | 4 | US-09-313-300-6 |
| 9 | 57 | 2.1 | 1129 | 4 | US-09-227-357-40 |
| 10 | 57 | 2.1 | 1325 | 1 | US-08-306-691B-51 |
| 11 | 56.6 | 2.0 | 1441 | 4 | US-08-821-994-63 |
| 12 | 56.4 | 2.0 | 1699 | 4 | US-09-152-060-19 |
| 13 | 56.4 | 2.0 | 2239 | 4 | US-09-196-390-1 |
| 14 | 56.2 | 2.0 | 577 | 4 | US-09-149-476-313 |
| 15 | 55.4 | 2.0 | 1325 | 2 | US-08-464-517-1 |
| 16 | 55.4 | 2.0 | 1325 | 2 | US-08-246-361A-1 |
| 17 | 55.4 | 2.0 | 1325 | 3 | US-08-463-772-1 |
| 18 | 55.4 | 2.0 | 1325 | 5 | PCT-US93-05000-1 |
| 19 | 55.2 | 2.0 | 2837 | 2 | US-08-993-228-11 |
| 20 | 54.4 | 2.0 | 1434 | 2 | US-09-055-097-2 |
| 21 | 54.2 | 2.0 | 458 | 1 | US-08-524-757-1 |
| 22 | 54.2 | 2.0 | 2246 | 4 | US-09-363-708-3 |
| 23 | 54.2 | 2.0 | 2447 | 2 | US-09-014-969-14 |
| 24 | 54.2 | 2.0 | 4843 | 3 | US-08-986-485-1 |
| 25 | 54 | 1.9 | 1825 | 4 | US-09-461-697-75 |
| 26 | 54 | 1.9 | 50000 | 4 | US-09-146-053-4 |
| 27 | 53.8 | 1.9 | 3214 | 1 | US-08-484-105-17 |

| | | | | | | |
|----|------|-----|------|---|-------------------|-------------------|
| 28 | 53.8 | 1.9 | 3214 | 1 | US-08-484-106-17 | Sequence 17, Appl |
| 29 | 53.8 | 1.9 | 3300 | 4 | US-09-336-643A-82 | Sequence 82, Appl |
| 30 | 53.8 | 1.9 | 6671 | 1 | US-08-280-443-1 | Sequence 1, Appl |
| 31 | 53.8 | 1.9 | 6671 | 1 | US-08-457-459-1 | Sequence 1, Appl |
| 32 | 53.8 | 1.9 | 6671 | 1 | US-08-555-678-1 | Sequence 1, Appl |
| 33 | 53.8 | 1.9 | 6671 | 5 | PCT-US95-02275-1 | Sequence 1, Appl |
| 34 | 53.4 | 1.9 | 1364 | 1 | US-08-265-087-3 | Sequence 3, Appl |
| 35 | 53.4 | 1.9 | 1364 | 1 | US-08-621-493-3 | Sequence 3, Appl |
| 36 | 53.4 | 1.9 | 1364 | 2 | US-08-965-688-3 | Sequence 3, Appl |
| 37 | 53.4 | 1.9 | 1364 | 4 | US-09-260-173-3 | Sequence 3, Appl |
| 38 | 53.4 | 1.9 | 2806 | 4 | US-09-653-839-9 | Sequence 9, Appl |
| 39 | 53.4 | 1.9 | 3080 | 4 | US-09-099-041A-25 | Sequence 25, Appl |
| 40 | 53.4 | 1.9 | 3080 | 4 | US-09-245-281-25 | Sequence 25, Appl |
| 41 | 53.4 | 1.9 | 3080 | 4 | US-09-207-359B-25 | Sequence 25, Appl |
| 42 | 53.2 | 1.9 | 1098 | 3 | US-09-248-335-35 | Sequence 35, Appl |
| 43 | 53.2 | 1.9 | 3410 | 4 | US-09-020-956-110 | Sequence 110, App |
| 44 | 53.2 | 1.9 | 3410 | 4 | US-09-030-607-110 | Sequence 110, App |
| 45 | 53.2 | 1.9 | 3410 | 4 | US-09-605-785-110 | Sequence 110, App |

ALIGNMENTS

RESULT 1
US-08-724-394A-16
; Sequence 16, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724.394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3502 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..3502
; OTHER INFORMATION: /note= "cdna 32"
US-08-724-394A-16

```
Query Match 4.8%; Score 133.6; DB 2; Length 3502;
Best Local Similarity 78.9%; Pred. No. 3.6e-27;
Matches 172; Conservative 0; Mismatches 44; Indels 2; Gaps 1;

QY 12 TTCTCTGCTGGAGCTGGGATATTCTTCTCTCTGCTGCTTGGACATCAGAGCTGCAGGC 71
Db 3101 TCTCTCTCTGAGAGCTGAGACACTCTTCTCTCTGCTGCTTGGACATCAGAACTCTCTGGC 3160

QY 72 TCTCTGGCTTTGGACCGGAGGATTTATACCAAGCAGGATTTCTGGGTCTTCAGGCCTTTG 131
Db 3161 TCTCTGGCTTTGAACTTCAAGACTTGTACCAGG--AGGCCCTGGGTCTTCAGGCCTTTG 3218

QY 132 GCCTTGGACTGATAGTTACACCATTTGGCATATCTGGTTCTGAGGCTCTTGGTCTTGGACT 191
Db 3219 GCTTTGGACTGAGATGTACAAATAGCTTCCCTGGTTCTGAGGCTTTCAGACTTAACT 3278

QY 192 GAGCCACACTCTGGCATCCAGCGCTTCCAGCTTGA 229
Db 3279 GAGCCATGCTACCAGCATCCAGGGTCTCCAGGCTACA 3316

RESULT 2
US-08-905-223-218
: Sequence 218, Application US/08905223
: Patent No. 6222029
: GENERAL INFORMATION:
: APPLICANT: Edwards, Jean-Baptiste D.
: APPLICANT: Duclert, Aymeric
: APPLICANT: Lacroix, Bruno
: TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
: NUMBER OF SEQUENCES: 503
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Knobbe, Martens, Olson & Bear
: STREET: 501 West Broadway
: CITY: San Diego
: STATE: California
: COUNTRY: USA
: ZIP: 92101-3505
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy Disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: Win95
: SOFTWARE: Word
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/905,223
: FILING DATE:
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Israel, Ned A.
: REGISTRATION NUMBER: 29,655
: REFERENCE/DOCKET NUMBER:
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 235-8550
: TELEFAX: (619) 235-0176
: INFORMATION FOR SEQ ID NO: 218:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 333 base pairs
: TYPE: NUCLEIC ACID
: STRANDEDNESS: DOUBLE
: TOPOLOGY: LINEAR
: MOLECULE TYPE: CDNA
: ORIGINAL SOURCE:
: ORGANISM: Homo Sapiens
: TISSUE TYPE: Brain
: FEATURE:
: NAME/KEY: other
: LOCATION: 32..331
: IDENTIFICATION METHOD: blastn
: OTHER INFORMATION: identity 99
: OTHER INFORMATION: region 1..300
: OTHER INFORMATION: id R13004
: OTHER INFORMATION: est

Query Match 4.5%; Score 124.8; DB 4; Length 333;
Best Local Similarity 74.2%; Pred. No. 2.5e-25;
Matches 184; Conservative 0; Mismatches 62; Indels 2; Gaps 2;

QY 24 AGCTGGATATTTCTTCTCTCTGCTTGGACATCAGAGCTGCAGGCTCTCTGGCCTTT 83
Db 2 AGCTAGGACATTTCTTCTCTCTGCTTGGACATCAGAACCCAGGCTCTCCAGCCTTT 61
QY 84 GGACCCGAGGATTTATACCAAGCAGGTTTCTGGGTCTTCAGGCTTTGGCCTTGA 143
```

```
Db 62 GGACTTCAGGACTGACA-CAAGAAACCTGCTGGGTTCTTAGGCTTTGG-CTTGACTGA 119
QY 144 TAGTTACACCATGGCATATCTGTTCTGAGGCTCTTGGTCTTGGACTGAGCCACACTCC 203
Db 120 GACTTACACCATCAGCTTCCCTGCTGAGACTTTTGGACTTGGATTTGAGCCACGCTAC 179
QY 204 TGCATCCCGAGCTCTCCAGCTTGGCATGGCTGCTGACGCTGATTCCTGTCAGAAATCACCAT 263
Db 180 TGCATCCCGAGGATCTCCAGCTTGCAGACAGCCTGCTGCTGGGACTTCACAGCCTCCATAA 239
QY 264 TTTTGGTA 271
Db 240 TTATAGAA 247
```

RESULT 3

```
US-09-482-180A-3
; Sequence 3, Application US/09482180A
; Patent No. 6361985
; GENERAL INFORMATION:
; APPLICANT: Konklin, Darrell C.
; APPLICANT: Yamamoto, Gayle
; APPLICANT: Gao, Zeren
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Jaspers, Stephen
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING
; TITLE OF INVENTION: BETA-1,3-GALACTOSYLTRANSFERASE HOMOLOG, ZNSSP6
; FILE REFERENCE: 98-80
; CURRENT APPLICATION NUMBER: US/09/482,180A
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/115,721
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: degenerate sequence
; LOCATION: (1)...(1134)
; OTHER INFORMATION: n is any nucleotide
; NAME/KEY: misc_feature
; LOCATION: (1)...(1134)
; OTHER INFORMATION: n = A,T,C or G
US-09-482-180A-3
```

```
Query Match 4.1%; Score 114.8; DB 4; Length 1134;
Best Local Similarity 31.3%; Pred. No. 3.3e-22;
Matches 209; Conservative 98; Mismatches 358; Indels 3; Gaps 1;

QY 504 TCCTTTGTTTACAGAAACACGGGAACTTCCCTTAAGCTCCAGATACAGACTCCAGGCAG 563
Db 286 TTYTNNACNTATGNCACAYTGYNAAATTTYSNATHYNTYNGARCCNWSNGNTGYWSN 345

QY 564 ACACCTCCCTCTCTGCTGCTGCTGCTATCCACAAACAGTTGCTGAGCGCATG 623
Db 346 AARGAYACNTTYTNTYNTGNCNATHAARWSNCARCCNGNCAYTNGARWNGNNGCN 405

QY 624 GCATCCCGCAGACGCTGGGGG---AAGAGAGGATGGTGAAGGAAAGAGCTGAAGACA 680
Db 406 GCNATHMGNWSNACNTGGGNGNGTNGGNGTGGCNGMNGNNGNCARYTNAARYTN 465

QY 681 TTCTTCTCTCTGGGACACAGCAGTGCAGCGGAAACAAAGAGGTGGACAGGAGAGC 740
Db 466 GTNTTYTNTYNTGNGTNGCNGNWSNCCNCCNCARYTNTYNTGNTAYGARWSN 525

QY 741 CAGCGACACGGGACATTTACAGAGGATTTCCCTAGAGCTCTATTACATCTGACCCCTG 800
Db 526 MGNARTTYGAYATHYNTTCARTGGGAYTTTACNGARGAYTTTYYTAAAYTNACNTYN 585
```

```
QY 801 AAGACCATGATGGGCATAGAAATGGTCCATCGCTTTTGTCTCCTCAGCGCGCTTGTGTGATG 860
Db 586 AARGARYTNCAYTNCARMGNTGGTNGTNGCNGCNTGYCCNCARGCNCAVTTATGYTN 645
QY 861 AAAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAAGTCTTCTGAAGAAA 920
Db 646 AARGNGAYGAYGAYGAYTNTTYTNCAYGTNCCNAAAYGTYNTGARTTYTNGAYGGNTGG 705
QY 921 AACAGAAACACCGAGTTTTCACCTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGCAG 980
Db 706 GAYCCNGCNCARGAYTNTYNTGNGNGAYGTNATHMNCARGONYTNCNAAATMGNAAY 765
QY 981 CCATTCAGCAAGTGGTTTGTTCAGTAAATCTCAATATCCGTGGGACAGTACCCACCATTC 1040
Db 766 ACNAARCTNAARTAYTTTATHCCNCCNWSNATGTAYWNGCNACNCAVATYCCNCNTAY 825
QY 1041 TGCTCCGCGCACCGCTACGTGTTTCTGGGACGCTGGGAGTCAAGTGTACATGTCCTCC 1100
Db 826 CGNGGNGGNGGNGTAYGTNATGWSNMGNCNACNGTNGMNGNYTNCARGCNATHATG 885
QY 1101 AAGAGCGTCCCATACATTAACCTGGAAGACGCTGTTGTGGGGCTCTGCCTCGAAAGGCTG 1160
Db 886 GARGAYCNGARYTNTYNTWSNATHGAYGTNTTYGTNGGNATGTGYTNGMNGNYTN 945
QY 1161 AACATCAG 1168
Db 946 GGNYTNS 953
```

RESULT 4

```
US-09-482-180A-1
; Sequence 1, Application US/09482180A
; Patent No. 6361985
; GENERAL INFORMATION:
; APPLICANT: Konklin, Darrell C.
; APPLICANT: Yamamoto, Gayle
; APPLICANT: Gao, Zeren
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Jaspers, Stephen
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING
; TITLE OF INVENTION: BETA-1,3-GALACTOSYLTRANSFERASE HOMOLOG, ZNSSP6
; FILE REFERENCE: 98-80
; CURRENT APPLICATION NUMBER: US/09/482,180A
; CURRENT FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/115,721
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1420
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (135)...(1271)
US-09-482-180A-1
```

```
Query Match 2.8%; Score 78.4; DB 4; Length 1420;
Best Local Similarity 46.1%; Pred. No. 6e-12;
Matches 299; Conservative 0; Mismatches 346; Indels 3; Gaps 1;
```

```
QY 524 CGGGAATTCCTTAAAGCTCCAGATACAGACTGCAGGCACACACCTCCCTTCCTCGCTCT 583
Db 440 CCAGAAATTTCTATCTTGTCTGGAGCTTCAGGCTGTTCAGAGATACCTTCTTGTCTCT 499

QY 584 GCTGTGACCTTCACACAAACAGTTGGCTGAGCGCATCGCCATCCGGCAGACGTGGGG 643
Db 500 GGCCATCAAGTACACGCTGCTCAGCTGGAGCGACGTCGCGCTATCCGACGACGTGGG 559

QY 644 GAAAGAGAGG---ATGGTGAAGGAAAGCAGCTGAAGACATTTCTCTCTCGGGGACAC 700
Db 560 CAGGGTGGGGGATGGCTAGGGGCGGCGAGCTGAAGCTGCTTCTCTCTAGGGGTGGC 619
```

| | | | |
|----|------|---|------|
| Qy | 701 | CAGCAGTGCAGCGGAACGAAGAGGTGGACAGAGAGACCGACACAGGGGACATTAT | 760 |
| Db | 620 | AGGATCCGCTCCCCACGCCACGCTCGCGCCTATGAGATAGGAGGTTTGATGACATCCT | 679 |
| Qy | 761 | CCAGAAGGATTTCCTAGACGCTATTACAATCTGACCTGAAGACCATGATGGGCATAGA | 820 |
| Db | 680 | CCAGTGGGACTTCAGTGAGACTTCTTCAACTGACGCTCAAGGAGCTGCACCTGCACGG | 739 |
| Qy | 821 | ATGGGTCCATCGCTTTTGTCTCTCAGCGCGGTTTGTGATGAAAACAGACTCAGACATGTT | 880 |
| Db | 740 | CTGGGTGGTGTGCTGCTGCCCGCAGGCCCATTTCTAGTCTAAGGAGATGACGATGTCCT | 799 |
| Qy | 881 | CATCAATGTGTGACTATCTCAGCTGAATGCTTCTGAAGAAAACAGACAACCGATTTT | 940 |
| Db | 800 | TGTCCAGCTCCCAACGTTGTAGATTCTGGATGGCTGGGACCCAGCCAGGACCTCCT | 859 |
| Qy | 941 | CAC TGGCTCTTTGAAACTCAATGAGTTTCCCATCAGGAGGCATTCAGCAAGTGGTTGT | 1000 |
| Db | 860 | GGTGGGAGATGTATCCGCCAAGCCCTGCCCAACAGGAACACTAAGGTCAAAATCTTCAT | 919 |
| Qy | 1001 | CAGTAAATCGAATATCCGTGGGACAGGTACCCACCATTCTGCTCCGCGACCCTACGT | 1060 |
| Db | 920 | CCACCCCTCAATGTACAGGGGCACCCACTACCACCCCTATGCTGGTGGGGGAGGATGT | 979 |
| Qy | 1061 | GTTTTCTGGCGAGCTGGCGAGTCAGGTGTACAAATGCTCTCCAAGAGCGTCCCATACATTA | 1120 |
| Db | 980 | CATGTCAGAGCCACAGTCGGGGCCCTCCAGGCTATCATGGAAGATGCTGAACCTCTCTC | 1039 |
| Qy | 1121 | ACTGGAAGACGTTGTTGTGGGGCTCTGGCTGGAAGGCTGAACATCAG | 1168 |
| Db | 1040 | CATTGATGATGCTTTGTGGGTATGTGCTCAGGAGGCTGGGGCTGAG | 1087 |

RESULT 5

```

US-09-459-133-3
; Sequence 3, Application US/09459133
; Patent No. 641898
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Yamamoto, Gayle
; APPLICANT: Yaspers, Stephen R.
; APPLICANT: Gao, Zeren
; TITLE OF INVENTION: BETA-1,3-GALACTOSYLTRANSFERASE HOMOLOGS

```

| | | | | |
|---------------------------|-------|------------------|------------|--------------|
| Query Match | 2.5% | Score 68.6; | DB 4; | Length 1191; |
| Best Local Similarity | 28.3% | Pred. No. 3e-09; | | |
| Matches 21l; Conservative | 95; | Mismatches 428; | Indels 12; | Gaps 3; |

[illegible]

| | | | | |
|----|------|------------|--|------|
| Qy | 690 | CTGGGGACCA | CCAGCAGCTGCAGCGGAAACGAAAGAGGTGGACGAGGAGCGACGACAC | 749 |
| | | | | |
| Db | 562 | GTNGNGARGC | NGCGNCGAYTYNGAYWS---NYTNGTNGCTTGGGARWSNMGNTAY | 618 |
| | | | | |
| Qy | 750 | GGGACATATT | CCAGAAGGATTCCTAGACGCTATTACAATCTGACCCCTCAAGACCATG | 809 |
| | | | | |
| Db | 619 | WSNGAYTYNT | NTNGGAYTTTTYNGAYTGCCNTTYAATCARACNTYTHAARGAYTN | 678 |
| | | | | |
| Qy | 810 | ATGGGCATAG | AATGGGTCCATCGCTTTTCTCTCAGCGGGCTTTGTGATGAAACACAGAC | 869 |
| | | | | |
| Db | 679 | YTYNTYNTG | CNTGGTYTNGNGMNCAYTCGCCNACNGTNSNTTYGTNTYTMNGMNCAR | 738 |
| | | | | |
| Qy | 870 | TCAGACATGT | TTCATCAATGTTGACTATCTGACTGAACGTCT-----TCTGAAGAAAAAC | 923 |
| | | | | |
| Db | 739 | GAYGAYCNT | TYTNCAYACNCCNGCYNTYTMNGNCAYTYTMNGMNCYTHCCNCCNGCN | 798 |
| | | | | |
| Qy | 924 | AGAACCAAC | CAGGTTTTTTCAGCTGGCTCTTTGAAACCAATGAGTTTCCCATCAGGCAGCA | 983 |
| | | | | |
| Db | 799 | WSNCGMNG | WSNYTYTAYTYTNGSGARGTNTTYACNARGCNGATGCCNTYTMGNAARCCN | 858 |
| | | | | |
| Qy | 984 | TTCAGCAAG | TGTGTTGTGACTAACTGAATATCCGTGGGACAGGTACCCACCATCTGCG | 1043 |
| | | | | |
| Db | 859 | GGNGGNCN | TTYAYTGTCNCCNARWSNTTYTYTGARGGNG---GNTAYCCNCGCNTAYGCN | 915 |
| | | | | |
| Qy | 1044 | TCGGGCACG | CGCTACCTGTTTTCTCGGACGTGGCGAGTCAAGTGTACAATGTCTCCAAG | 1103 |
| | | | | |
| Db | 916 | WSNGNGNG | NGNTAYGTNATHGCGNGMNYTMNGNCCNTGGTYNTYTMNGMNGCNGCNGCN | 975 |
| | | | | |
| Qy | 1104 | AGCGTCCCA | TACATTAACCTGGAACACGTGTTGTGGGGCTCTGCTCGAAAGGCTGAAC | 1163 |
| | | | | |
| Db | 976 | MNGTNGCN | CNTTYCNGTYTGARGAYGTNAYACNGGNYNTGYATHMGNGCYTMNGCN | 1035 |
| | | | | |
| Qy | 1164 | ATCAGATTG | GAGAGCTCCACTCCACGCGGACCTTTTTTCCAGGGGGCTTACGCTTCTCC | 1223 |
| | | | | |
| Db | 1036 | YTAGTNGCN | CARGCNCAYCCNGGNTTYTNACNGTNGTGGCCNGCNGAYTMGNACNGCNGAY | 1095 |
| | | | | |
| Qy | 1224 | GTAATGCTC | TCTTCAGGAGGATCGTGGCCTGCCACTTCATCAAGCCCTCGGACCTCTTGAC | 1283 |
| | | | | |
| Db | 1096 | CAYTGYCN | TTYMGNAAYYTYNTYTMNGTMNGMNCNYTMNGMNCNARGCWSNATHMGN | 1155 |
| | | | | |
| Qy | 1284 | TACTGGCAG | CGCTCTAGAGAAATCCCG | 1309 |
| | | | | |
| Db | 1156 | YTNITGGAR | CARYTNCARGAYCCNMG | 1181 |

```

RESULT 6
US-09-459-133-14
; Sequence 14, Application US/09459133
; Patent No. 6416988
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Yamamoto, Gayle
; APPLICANT: Jaspers, Stephen R.
; APPLICANT: Gao, Zeren
; TITLE OF INVENTION: BETA-1,3-GALACTOSYLTRANSFERASE HOMOLOGS
; FILE REFERENCE: 98-77
; CURRENT APPLICATION NUMBER: US/09/459,133
; CURRENT FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/111,697
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 1167
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: degenerate sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1167)
; OTHER INFORMATION: n = A,T,C or G

```


| Db | 1013 | AAATGGTGAGCAATGAAGACAATAAAATTTTTCATTTCCNAAAAAAAAAAAAAAAAAAAAA | 1072 |
|--|------|---|------|
| Qy | 2742 | AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA | 2775 |
| Db | 1073 | AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA | 1106 |
| RESULT 10 | | | |
| US-08-306-691B-51 | | | |
| ; Sequence 51, Application US/08306691B | | | |
| ; Patent No. 5734039 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: Calabretta, Bruno | | | |
| ; APPLICANT: Skorski, Tomasz | | | |
| ; TITLE OF INVENTION: ANTISENSE | | | |
| ; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES | | | |
| ; NUMBER OF SEQUENCES: 55 | | | |
| ; CORRESPONDENCE ADDRESS: | | | |
| ; ADDRESSEE: Seidel, Gonda, Lavorigna & Monaco, P.C. | | | |
| ; STREET: Two Penn Center, Suite 1800 | | | |
| ; CITY: Philadelphia | | | |
| ; STATE: Pennsylvania | | | |
| ; COUNTRY: U.S.A. | | | |
| ; ZIP: 19102 | | | |
| ; COMPUTER READABLE FORM: | | | |
| ; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb | | | |
| ; COMPUTER: IBM PS/2 | | | |
| ; OPERATING SYSTEM: MS-DOS | | | |
| ; SOFTWARE: WordPerfect 5.1 | | | |
| ; CURRENT APPLICATION DATA: | | | |
| ; APPLICATION NUMBER: US/08/306,691B | | | |
| ; FILING DATE: September 15, 1994 | | | |
| ; CLASSIFICATION: 514 | | | |
| ; PRIOR APPLICATION DATA: | | | |
| ; APPLICATION NUMBER: | | | |
| ; FILING DATE: | | | |
| ; ATTORNEY/AGENT INFORMATION: | | | |
| ; NAME: Monaco, Daniel A. | | | |
| ; REGISTRATION NUMBER: 30,480 | | | |
| ; REFERENCE/DOCKET NUMBER: 8321-8 | | | |
| ; TELECOMMUNICATION INFORMATION: | | | |
| ; TELEPHONE: (215) 568-8383 | | | |
| ; TELEX: (215) 568-5549 | | | |
| ; TELEFAX: (215) 568-5549 | | | |
| ; INFORMATION FOR SEQ ID NO: 51: | | | |
| ; SEQUENCE CHARACTERISTICS: | | | |
| ; LENGTH: 1325 base pairs | | | |
| ; TYPE: nucleic acid | | | |
| ; STRANDEDNESS: single | | | |
| ; TOPOLOGY: linear | | | |
| US-08-306-691B-51 | | | |
| Query Match 2.1%; Score 57; DB 1; Length 1325; | | | |
| Best Local Similarity 62.1%; Pred. NO. 5.7e-06; | | | |
| Matches 90; Conservative 0; Mismatches 55; Indels 0; Gaps | | | |
| Qy | 2631 | TGATTTTCTTTTAAATTTTTTACTTTACCAGACTTTACTTTGTACTCAGACAAGAGCCT | 2690 |
| Db | 1167 | TGTTTTTCTTTTGTCTTTTCCCTTCCATCTCTGACTTAAAGCAAAAGAAAAAGATTAC | 1226 |
| Qy | 2691 | CACATGGCTGTGCACATATAAATGTGGACTAAACTCTTAAAAA | 2750 |
| Db | 1227 | CCAAACATGCTCTTTAAAGAGAGAGAGAAAAA | 1286 |
| Qy | 2751 | AAAAAAAAAAAAAAAAAAAAA | 2775 |
| Db | 1287 | AAAAAAAAAAAAAAAAAAAAA | 1311 |
| RESULT 11 | | | |
| US-08-821-994-63 | | | |
| ; Sequence 63, Application US/08821994A | | | |

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (871)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-152-060-19

Query Match          2.0%; Score 56.4; DB 4; Length 1699;
Best Local Similarity 80.5%; Pred. No. 9,8e-06;
Matches 66; Conservative 0; Mismatches 16; Indels 0; Gaps
;

QY 2694 ATGGCTGTGCACATATAAATGTTGGACTAAACTCTTAAAAA 2753
      ||||| | ||| ||| ||| | ||| ||||| ||||| |||||
Db 1610 ATGCCTTCGTAATAAAAAAGATTGGGATTTCTTTTGAAAAA 1669
      ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 2754 AAAAAAAAAAAAAAAAAAAAAA 2775
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1670 AAAAAAAAAAAAAAAAAAAAAA 1691
      ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
US-09-196-390-1
; Sequence 1, Application US/09196390
; Patent No. 6307125
; GENERAL INFORMATION:
; APPLICANT: Block, Martina
; APPLICANT: Lorz, Horst
; APPLICANT: Lutticke, Stephanie
; APPLICANT: Walter, Lennart
; APPLICANT: Frohberg, Claus
; APPLICANT: Kossmann, Jens
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
; TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH
; TITLE OF INVENTION: SYNTHESIS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196,390
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 21 588.9
; FILING DATE: 29-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 36 917.7
; FILING DATE: 11-SEP-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP97/02793
; FILING DATE: 28-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: AGREVO-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2239 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

```
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Triticum aestivum L.
; STRAIN: cv. Florida
; HAPLOTYPE: ca. 21 d Caryopses
; IMMEDIATE SOURCE:
; LIBRARY: cDNA library in pluescript sk (-)
; CLONE: TaSSS
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..2017
; US-09-196-390-1

Query Match 2.0%; Score 56.4; DB 4; Length 2239;
Best Local Similarity 63.0%; Pred. No. 1.2e-05;
Matches 87; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 2638 TCTTTTAATTTTCTTACTTTTACCAGACTTTTCTGCTCAGAGAGGCGCTCACATGG 2697
      ||| || ||||| || || || || || || || || || || || || || || || ||
Db 2088 TGTTCGAGAATTTTATCTTGTCTGCTGTTTTTTTTTAAATCAAAAGAGGGGTTTCTCCG 2147

QY 2698 CTGCTGCACATATAATGTTGGACTTAAACTCTTAAAAAAGAGAGAGAGAGAGAGAGAGAG 2757
      ||| || || || || || || || || || || || || || || || || || || ||
Db 2148 ATTCATTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2207

QY 2758 AAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2775
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2208 AAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2225

RESULT 14
US-09-149-476-313
; Sequence 313, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618

; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
```


SEQUENCE CHARACTERISTICS:
LENGTH: 1325 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 145..1029
US-08-464-517-1

Query Match 2.0%; Score 55.4; DB 2; Length 1325;
Best Local Similarity 61.4%; Pred. No. 1.6e-05;
Matches 89; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY 2631 TGATTTTCTTTTAAATTTTACTTTTACCAGACTTTTACTTTGTGACTCAGAGAAGAGCCCT 2690
DB 1167 TGTCTTTCTTTGCTCTTTCCCTTTCCATCTCTCACTTAACCAACAAAAAGATTAC 1226
QY 2691 CACATGGCTGTGCACATATAATTTGGACTAACTCTTAAAAA
DB 1227 CAAAACCTGCTTTAAAGAGAGAGAGAAAAA
QY 2751 AAAAAA
DB 1287 AAAAAA

Search completed: April 11, 2003, 22:41:26
Job time : 111.627 secs

Db 337 TAGCCCAAGAGGCTGAAAGGAAGACAAAAATGTTGGCAATTTCTCTAGAGGATGAACACC 396
QY 742 AGGACACAGGGGACATATCCAGAGGATTTCTTAGACGCTCTATTACAATCTGACCCCTGA 801
Db 397 TTCTTTATGTTGACATTAATCCGACAGATTTTTTAGACACATATAATAAACCTGACCTTGA 456
QY 802 AGACCATGATGGGCATAGAAATGGGTCCATCGCTTTTCTCAGCGCGGCTTTGTGATGA 861
Db 457 AAACCATATGGCATTCAGGTGGTAACTGAGTTTTCGCCCAATGCCAAGTACGTAATGA 516
QY 862 AAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAATCTCTCTGAGAAAA 921
Db 517 AGACAGACACTGATGTTTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAA 576
QY 922 ACAGAACCAACAGGTTTTTCACTGGCTTCTTGAACATCAATGAGTTTCCCATCAGCGCAGC 981
Db 577 ACCACTCAGAGAAAGTTTTTTCACAGGTTATCTCTAATTTGATAATTTCTCTATAGAGAT 636
QY 982 CATTGACCAAGTGTGTTGTCAGTAATACTGAATATCCGTGGGACAGGTACCCACCATCT 1041
Db 637 TTTACCAAAAACCCATATTTCTTACCAGGAGTATCTTTTCAAGGTGTTCCCTCCATACT 696
QY 1042 GCTCCGCGCAGGCTAGCTGTTTCTGCGGAGTGGCGAGTCAAGTGTACAAATGCTCTCCA 1101
Db 697 GCAGTGGGTTGGTTATATAATGTCACAGATTTGGTGCCCAAGGATCTATGAATGATGG 756
QY 1102 AGAGCGTCCCATCAATTAATACTGGAAGACGCTGTTTGTGGGCTCTGCTCGAAAGGCTGA 1161
Db 757 GTCACGTAACCAACCATCAAGTTTGAAGATGTTTATGTCGGGATCTGTTGAATTTATTAA 816
QY 1162 ACATCAGATTTGGAGGAGCTCCACTCCAGCGGACCTTTTTTCCAGGGGGCTTAGCGCTCT 1221
Db 817 AAGTGAACATTCATATTTCCAGAGACACAAATCTTTTCTTATATAGAAATCCATTGG 876
QY 1222 CCGTATCGCTTCTCAGGAGGATCGTGGCCTGCCACTTCATCAAGCCCTCGGACTCTCTGG 1281
Db 877 ATGTCGTCAACTGAGAGCGTGTGATTCCAGCCCATGCTTTCTTCCAGGAGATCATCA 936
QY 1282 ACTACTGGCAGG 1293
Db 937 CTTTGGCAGG 948

RESULT 2

US-09-992-598-208

Sequence 208, Application US/09992598

Patent No. US20020160384A1

GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.

APPLICANT: Baker, Kevin P.

APPLICANT: Botstein, David

APPLICANT: Desnovers, Luc

APPLICANT: Eaton, Dan L.

APPLICANT: Ferrara, Napoleone

APPLICANT: Fong, Sherman

APPLICANT: Gerber, Hanspeter

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, J. Christopher

APPLICANT: Gurney, Austin L.

APPLICANT: Kljavin, Ivar J.

APPLICANT: Napier, Mary A.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K.

APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

FILE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730P1C20
CURRENT APPLICATION NUMBER: US/09/992,598
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861

| | | |
|----|-------------------------------------|--|
| 1 | PRIOR FILING DATE: 1998-06-11 | |
| 2 | PRIOR APPLICATION NUMBER: 60/088876 | |
| 3 | PRIOR FILING DATE: 1998-06-11 | |
| 4 | PRIOR APPLICATION NUMBER: 60/089105 | |
| 5 | PRIOR FILING DATE: 1998-06-12 | |
| 6 | PRIOR APPLICATION NUMBER: 60/089440 | |
| 7 | PRIOR FILING DATE: 1998-06-16 | |
| 8 | PRIOR APPLICATION NUMBER: 60/089512 | |
| 9 | PRIOR FILING DATE: 1998-06-16 | |
| 10 | PRIOR APPLICATION NUMBER: 60/089514 | |
| 11 | PRIOR FILING DATE: 1998-06-16 | |
| 12 | PRIOR APPLICATION NUMBER: 60/089532 | |
| 13 | PRIOR FILING DATE: 1998-06-17 | |
| 14 | PRIOR APPLICATION NUMBER: 60/089538 | |
| 15 | PRIOR FILING DATE: 1998-06-17 | |
| 16 | PRIOR APPLICATION NUMBER: 60/089598 | |
| 17 | PRIOR FILING DATE: 1998-06-17 | |
| 18 | PRIOR APPLICATION NUMBER: 60/089599 | |
| 19 | PRIOR FILING DATE: 1998-06-17 | |
| 20 | PRIOR APPLICATION NUMBER: 60/089600 | |
| 21 | PRIOR FILING DATE: 1998-06-17 | |
| 22 | PRIOR APPLICATION NUMBER: 60/089653 | |
| 23 | PRIOR FILING DATE: 1998-06-17 | |
| 24 | PRIOR APPLICATION NUMBER: 60/089801 | |
| 25 | PRIOR FILING DATE: 1998-06-18 | |
| 26 | PRIOR APPLICATION NUMBER: 60/089907 | |
| 27 | PRIOR FILING DATE: 1998-06-18 | |
| 28 | PRIOR APPLICATION NUMBER: 60/089908 | |
| 29 | PRIOR FILING DATE: 1998-06-18 | |
| 30 | PRIOR APPLICATION NUMBER: 60/089947 | |
| 31 | PRIOR FILING DATE: 1998-06-19 | |
| 32 | PRIOR APPLICATION NUMBER: 60/089948 | |
| 33 | PRIOR FILING DATE: 1998-06-19 | |
| 34 | PRIOR APPLICATION NUMBER: 60/089952 | |
| 35 | PRIOR FILING DATE: 1998-06-19 | |
| 36 | PRIOR APPLICATION NUMBER: 60/090246 | |
| 37 | PRIOR FILING DATE: 1998-06-22 | |
| 38 | PRIOR APPLICATION NUMBER: 60/090252 | |
| 39 | PRIOR FILING DATE: 1998-06-22 | |
| 40 | PRIOR APPLICATION NUMBER: 60/090254 | |
| 41 | PRIOR FILING DATE: 1998-06-22 | |
| 42 | PRIOR APPLICATION NUMBER: 60/090349 | |
| 43 | PRIOR FILING DATE: 1998-06-23 | |
| 44 | PRIOR APPLICATION NUMBER: 60/090355 | |
| 45 | PRIOR FILING DATE: 1998-06-23 | |
| 46 | PRIOR APPLICATION NUMBER: 60/090429 | |
| 47 | PRIOR FILING DATE: 1998-06-24 | |
| 48 | PRIOR APPLICATION NUMBER: 60/090431 | |
| 49 | PRIOR FILING DATE: 1998-06-24 | |
| 50 | PRIOR APPLICATION NUMBER: 60/090435 | |
| 51 | PRIOR FILING DATE: 1998-06-24 | |
| 52 | PRIOR APPLICATION NUMBER: 60/090444 | |
| 53 | PRIOR FILING DATE: 1998-06-24 | |
| 54 | PRIOR APPLICATION NUMBER: 60/090445 | |
| 55 | PRIOR FILING DATE: 1998-06-24 | |
| 56 | PRIOR APPLICATION NUMBER: 60/090472 | |
| 57 | PRIOR FILING DATE: 1998-06-24 | |
| 58 | PRIOR APPLICATION NUMBER: 60/090535 | |
| 59 | PRIOR FILING DATE: 1998-06-24 | |
| 60 | PRIOR APPLICATION NUMBER: 60/090540 | |
| 61 | PRIOR FILING DATE: 1998-06-24 | |
| 62 | PRIOR APPLICATION NUMBER: 60/090542 | |
| 63 | PRIOR FILING DATE: 1998-06-24 | |
| 64 | PRIOR APPLICATION NUMBER: 60/090557 | |
| 65 | PRIOR FILING DATE: 1998-06-24 | |
| 66 | PRIOR APPLICATION NUMBER: 60/090676 | |
| 67 | PRIOR FILING DATE: 1998-06-25 | |
| 68 | PRIOR APPLICATION NUMBER: 60/090678 | |
| 69 | PRIOR FILING DATE: 1998-06-25 | |
| 70 | PRIOR APPLICATION NUMBER: 60/090690 | |
| 71 | PRIOR FILING DATE: 1998-06-25 | |
| 72 | PRIOR APPLICATION NUMBER: 60/090694 | |
| 73 | PRIOR FILING DATE: 1998-06-25 | |

1 PRIOR APPLICATION NUMBER: 60/090695
 2 PRIOR FILING DATE: 1998-06-25
 3 PRIOR APPLICATION NUMBER: 60/090696
 4 PRIOR FILING DATE: 1998-06-25
 5 PRIOR APPLICATION NUMBER: 60/090862
 6 PRIOR FILING DATE: 1998-06-26
 7 PRIOR APPLICATION NUMBER: 60/090863
 8 PRIOR FILING DATE: 1998-06-26
 9 PRIOR APPLICATION NUMBER: 60/091360
 10 PRIOR FILING DATE: 1998-07-01
 11 PRIOR APPLICATION NUMBER: 60/091478
 12 PRIOR FILING DATE: 1998-07-02
 13 PRIOR APPLICATION NUMBER: 60/091544
 14 PRIOR FILING DATE: 1998-07-01
 15 PRIOR APPLICATION NUMBER: 60/091519
 16 PRIOR FILING DATE: 1998-07-02
 17 PRIOR APPLICATION NUMBER: 60/091626
 18 PRIOR FILING DATE: 1998-07-02
 19 PRIOR APPLICATION NUMBER: 60/091633
 20 PRIOR FILING DATE: 1998-07-02
 21 PRIOR APPLICATION NUMBER: 60/091978
 22 PRIOR FILING DATE: 1998-07-07
 23 PRIOR APPLICATION NUMBER: 60/091982
 24 PRIOR FILING DATE: 1998-07-07
 25 PRIOR APPLICATION NUMBER: 60/092182
 26 PRIOR FILING DATE: 1998-07-09

Query Match 4.7%; Score 130.4; DB 9; Length 2095;
Best Local Similarity 48.9%; Pred. No. 2.8e-25;
Matches 387; Conservative 0; Mismatches 396; Indels 9;

| | | | |
|----|------|--|------|
| Qy | 511 | TTTACAAGAAAGCGGGAACCTTCCTTAAGCTCCAGATACAGACTGCAGGCAGACACCTC | 570 |
| Db | 494 | TTTTACAGAACAAGCTTTCACTTCACACTTCGAGAGCATTCAAACCTGCTCTCACTAAAATC | 553 |
| Qy | 571 | CCTTCCTGCTCGTGCTGACCTCATCCCACAAACAGTTGGGTGAGCGCATGGCCATCC | 630 |
| Db | 554 | CATTCTGGTCACTCTGGTGACCTCCCACCCTTCAGATGTAAAGCCAGCGAGGCATTA | 613 |
| Qy | 631 | GGCAGAGCTGGGAAAGAGAGGATGGTGAAGGAAAGCAGCTGAAGACATTTCTTCCTCC | 690 |
| Db | 614 | GAGTTACTTGGGTGAAAAAAGCTTTGGTGGGATATGAGGTCTCTTACATTTTCTTAT | 673 |
| Qy | 691 | TGGGACCACACAGCAGTGCAGCGGACGAACGAAGAGG-----TGGACCAAGGAGGCC | 741 |
| Db | 674 | TAGCCCAAGAGGCTGAAAAGGAAGACAAATGTTGGCATGTCTTAGAGGATGAACACC | 733 |
| Qy | 742 | AGCGACAGGGGACATATTCCAGAAGGATTTCTTAGACGCTCTATTACAATCTGACCCCTGA | 801 |
| Db | 734 | TTCTTTATGCTGACATAATCCGACAAGATTTTTTAGACACATATAATAACCTGACCTTGA | 793 |
| Qy | 802 | AGACCATGATGGCCATAGAATGGTCCATCGCTTTTGTCCTCAGGCGGCGTTGTGATGA | 861 |
| Db | 794 | AACCATATTGGCATTCAGGTGGGTAACTGAGTTTTTGCCCCCAATGCCAAGTACGTAATGA | 853 |
| Qy | 862 | AAACAGACTCAGACATGTTTCATCAATGTTTGACTATCTGACTGAACTGCTTCTTGAAGAAA | 921 |
| Db | 854 | AGACAGACATGATGTTTTTCATCAATACTGCGCAATTTAGTGAAGTATCTTTTAAACCTAA | 913 |
| Qy | 922 | ACAGAACAACACGTTTTTTCACCTGGCTCTTGAAACTCAATCAGTTTCCCATCAGGCAGC | 981 |
| Db | 914 | ACCACCTCAGAGAAGTTTTTCACAGGTTATCTCTTAATTGATTAATTTCCTATAGAGAT | 973 |
| Qy | 982 | CATTACAGAAAGTGGTTGTTCAGTAAATCTGAATATCCGTGGACAGGTACCCACCATCT | 1041 |
| Db | 974 | TTTACCAAAAAAACCCATATTTCTTACAGGAGTATCTTTCAAGGTGTTCCTCCCTACT | 1033 |
| Qy | 1042 | GCTCCGCGCACGGCTACGTGTTTTCTTGGCGAGTGGCGAGTCAGGTGACATGTCCTCA | 1101 |
| Db | 1034 | GCAGTGGGTGGGTATATAATGTCCAGAGATTTGGTCCCAAGGATCTATGAATGATGG | 1093 |
| Qy | 1102 | AGAGCTGCCATACATAAACTGGAAGCGTGTTTGTGGGGCTCTGCCGTAAGAGGCTGA | 1161 |

Db 1094 GTCACGTAACCCATCAAGTTTGAAGATGTTTATGTCGGGATCTGTTTGAATTTATTAA 1153
QY 1162 ACATCAGATTGGAGGAGCTCCACTCCAGCCGACCTTTTCCAGGGGCTTACGCTTCT 1221
Db 1154 AAGTGAACATTCATATTCAGAAGACACAATCTTCTTCTTATATAGAAATCCATTGG 1213
QY 1222 CCATATGCCCTCTTCAGGAGGATCGTGGCCCTGGCCACTTCATCAAGCCCTCGGACTCTCTTGG 1281
Db 1214 ATCTCTGTCACATGAGAGCGTGTGATTGCAGCCCATGCGCTTTCTTCCAGGAGATCATCA 1273
QY 1282 ACTACTGGCAGG 1293
Db 1274 CTTTGGCAGG 1285

RESULT 3

US-09-989-293A-208

; Sequence 208, Application US/09989293A

; Patent No. US20020177164A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James

; APPLICANT: Faonl, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: P2730P1C66

; CURRENT APPLICATION NUMBER: US/09/989,293A

; CURRENT FILING DATE: 2001-11-20

; PRIOR APPLICATION NUMBER: 60/049787

; PRIOR FILING DATE: 1997-06-16

; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/065186

; PRIOR FILING DATE: 1997-11-12

; PRIOR APPLICATION NUMBER: 60/065311

; PRIOR FILING DATE: 1997-11-13

; PRIOR APPLICATION NUMBER: 60/066770

; PRIOR FILING DATE: 1997-11-24

; PRIOR APPLICATION NUMBER: 60/075945

; PRIOR FILING DATE: 1998-02-25

; PRIOR APPLICATION NUMBER: 60/078910

; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/083322

; PRIOR FILING DATE: 1998-04-28

; PRIOR APPLICATION NUMBER: 60/084600

; PRIOR FILING DATE: 1998-05-07

; PRIOR APPLICATION NUMBER: 60/087106

; PRIOR FILING DATE: 1998-05-28

; PRIOR APPLICATION NUMBER: 60/087607

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/087609

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/089908

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/089908

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/089908

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/089908

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/089908

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908

| | | | | | |
|----|------|-----------------------|---------------------|---------------------------------|------|
| Qy | 511 | TTTACAGAAAGACGGGAAC | TCTCCTTTAAGCTCC | CAGATACAGACTGCAGGACAGACCTC | 570 |
| Db | 494 | TTTACAGCAAGAACTTTCACT | TTTCACACTTCGAGAGCA | TTCAAACTGCTCTCATCAAAATC | 553 |
| Qy | 571 | CTTTCTCGTCGCTGGTGAC | CTCATCCCAACAGTTGGCT | CTGAGCGCATGGCCATCC | 630 |
| Db | 554 | CATTTCTGGTCACTTCGGT | GACCTCCCAACCCTTCAG | ATGTGAAGCCAGGCGAGCCATTA | 613 |
| Qy | 631 | GGCAGACGTGGGGGAAGAG | AGGATGGTGAAGGGGAAG | CAGCTGAAGCATTTCTTCCTCC | 690 |
| Db | 614 | GAGTTACTTGGGTGAAAAA | AGCTCTTGCTGGGATATC | AGGTTCTTACATTTTCTTAT | 673 |
| Qy | 691 | TGGGGACCAACGACGTGC | AGCGGGAACGAAGAGG | -----TGGACCAAGGAGACC | 741 |
| Db | 674 | TAGGCCAAGAGGCTGAAAA | AGGAACAAATGTTGGCAT | TGTTCCTTAGAGGATGAACACC | 733 |
| Qy | 742 | AGCGACACGGGACATATTCC | AGAAGATTTCCCTAGACG | CTTATTACAACTCGACCTCGA | 801 |
| Db | 734 | TTCTTATGCTGACATAATCC | GACAAGATTTTTTAGCAC | ACATATAATRACCTTGACCTTGA | 793 |
| Qy | 802 | AGACACTGCAGCATGTTTC | ATCAATGTTTGACTATCT | GACTGAACTGCTTCTGGAAGAAA | 921 |
| Db | 794 | AAACCATTTATGGCATTC | AGTGAGTTTGCCCAATG | CCCAAGTACGTAATGA | 853 |
| Qy | 862 | AAACAGACTGCAGCATGTT | TCATCAATGTTTGACTAT | CTGACTGAACTGCTTCTGGAAGAAA | 921 |
| Db | 854 | AGACAGACACTGATGTTTT | TCATCAATGCTGCAAT | TTAGTGAAGTATCTTTTAAACCTAA | 913 |
| Qy | 922 | ACAGAACAAACAGGTTTT | TTTCACTGGCTCTTGAA | ACTCAATCAGTTTCCCATCAGGCAGC | 981 |
| Db | 914 | ACCACCTCAGAGATTTTT | TCACAGGTTATCTCT | TAATGATAATTTATTCCTATAGAGAT | 973 |
| Qy | 982 | CATTACAGCAAGTGGTTT | GTGAGTAATCTCGATG | AGACAGTACCCGACAGGTACCCACCATTTCT | 1041 |
| Db | 974 | TTTACCAAAAAACCCAT | ATTTCTTACCAAGGAT | ATCCTTTCAGGTGTTCCCTCCATACT | 1033 |
| Qy | 1042 | GCCTCCGCACCGCTACG | TGTTTTCTGCGCAGTGG | CAGTGGCAGTGTACAACTGTCTCCA | 1101 |
| Db | 1034 | GCAGTGGGTGGGTTAT | ATAATGTCAGAGATTT | GGTGCCCAAGGATCTATGAAATGATGG | 1093 |
| Qy | 1102 | AGAGCGTCCCATACATTA | AACTGGAAGACGCTGTT | GTGGGGCTCTGCCTCGAAAGGCTGA | 1161 |
| Db | 1094 | GTCAGCTAAACCCCATCA | AGTTTGAAGATGTTAT | GTGCGGATCTGTTTGAATTTATTAA | 1153 |
| Qy | 1162 | ACATCAGATTGGAGAGCT | CTCCACTCCAGCGACG | CTTTTTTCCAGGGGGCTTACGCTTCT | 1221 |
| Db | 1154 | AAGTGAAACATTCATAT | TCCAGAAAGACAAAA | CTTTTCTCTATATAGAATCCATTTGG | 1213 |
| Qy | 1222 | CGTATGCTCTCTCAGG | AGGATCGTGGCTCGCA | CTTTCATCAAGCCTCGGACTCTCTTGG | 1281 |
| Db | 1214 | ATGCTGTCAACTGAGAC | GCTGTGATGTCAGCC | CTTGTTCCTCCAAGGAGATCATCA | 1273 |
| Qy | 1282 | ACTACTGGCGAGG | 1293 | | |
| Db | 1274 | CTTTTGGCGAGG | 1285 | | |

| | |
|---|--------------------------------------|
| ? | PRIOR APPLICATION NUMBER: 60/0903419 |
| ? | PRIOR FILING DATE: 1998-06-23 |
| ? | PRIOR APPLICATION NUMBER: 60/090355 |
| ? | PRIOR FILING DATE: 1998-06-23 |
| ? | PRIOR APPLICATION NUMBER: 60/090429 |
| ? | PRIOR FILING DATE: 1998-06-24 |
| ? | PRIOR APPLICATION NUMBER: 60/090431 |
| ? | PRIOR FILING DATE: 1998-06-24 |
| ? | PRIOR APPLICATION NUMBER: 60/090435 |
| ? | PRIOR FILING DATE: 1998-06-24 |
| ? | PRIOR APPLICATION NUMBER: 60/090444 |
| ? | PRIOR FILING DATE: 1998-06-24 |
| ? | PRIOR APPLICATION NUMBER: 60/090445 |
| ? | PRIOR FILING DATE: 1998-06-24 |
| ? | PRIOR APPLICATION NUMBER: 60/090472 |
| ? | PRIOR FILING DATE: 1998-06-24 |
| ? | PRIOR APPLICATION NUMBER: 60/090535 |
| ? | PRIOR FILING DATE: 1998-06-24 |
| ? | PRIOR APPLICATION NUMBER: 60/090540 |
| ? | PRIOR FILING DATE: 1998-06-24 |
| ? | PRIOR APPLICATION NUMBER: 60/090542 |
| ? | PRIOR FILING DATE: 1998-06-24 |
| ? | PRIOR APPLICATION NUMBER: 60/090557 |
| ? | PRIOR FILING DATE: 1998-06-24 |
| ? | PRIOR APPLICATION NUMBER: 60/090676 |
| ? | PRIOR FILING DATE: 1998-06-25 |
| ? | PRIOR APPLICATION NUMBER: 60/090678 |
| ? | PRIOR FILING DATE: 1998-06-25 |
| ? | PRIOR APPLICATION NUMBER: 60/090690 |
| ? | PRIOR FILING DATE: 1998-06-25 |
| ? | PRIOR APPLICATION NUMBER: 60/090694 |
| ? | PRIOR FILING DATE: 1998-06-25 |
| ? | PRIOR APPLICATION NUMBER: 60/090695 |
| ? | PRIOR FILING DATE: 1998-06-25 |
| ? | PRIOR APPLICATION NUMBER: 60/090696 |
| ? | PRIOR FILING DATE: 1998-06-25 |
| ? | PRIOR APPLICATION NUMBER: 60/091360 |
| ? | PRIOR FILING DATE: 1998-07-01 |
| ? | PRIOR APPLICATION NUMBER: 60/091478 |
| ? | PRIOR FILING DATE: 1998-07-02 |
| ? | PRIOR APPLICATION NUMBER: 60/091544 |
| ? | PRIOR FILING DATE: 1998-07-01 |
| ? | PRIOR APPLICATION NUMBER: 60/091519 |
| ? | PRIOR FILING DATE: 1998-07-02 |
| ? | PRIOR APPLICATION NUMBER: 60/091626 |
| ? | PRIOR FILING DATE: 1998-07-02 |
| ? | PRIOR APPLICATION NUMBER: 60/091633 |
| ? | PRIOR FILING DATE: 1998-07-02 |
| ? | PRIOR APPLICATION NUMBER: 60/091978 |
| ? | PRIOR FILING DATE: 1998-07-07 |
| ? | PRIOR APPLICATION NUMBER: 60/091982 |
| ? | PRIOR FILING DATE: 1998-07-07 |
| ? | PRIOR APPLICATION NUMBER: 60/092182 |
| ? | PRIOR FILING DATE: 1998-07-09 |

Matches 387; **Conservative** 0; **Mismatches** 396; **Indels** 9; **Gaps** 1;

| | | | |
|----|-----|--|-----|
| Qy | 511 | TTTACAAGAAAGACGGAACTTCTTAAGTCCCATACAGACTCGAGGCAGACACCTC | 570 |
| | | | |
| | | | |
| | | | |
| | | | |
| Db | 494 | TTTACAGACAGACTTTCATCTTCACACTTCGAGAGCATTCAAACTGCTCTCATCAAAATC | 553 |
| | | | |
| | | | |
| | | | |
| | | | |
| Qy | 571 | CTTTCTCTGCTCTGCTGGTGACCTTCATCCACAAACAGTTGGCTGAGCGCATGCCATCC | 630 |
| | | | |
| | | | |
| | | | |
| Db | 554 | CATTCTGGTCATCTCTGCTGACCTCCCAACCTTCAGATGTGAAGCCAGCGAGGCCATT | 613 |
| | | | |
| | | | |
| | | | |
| Qy | 631 | GGCAGCGTGGGGAAAGAGAGGATGTTGAAGGGAAAGCAGCTGAAGACATTTCTTCCTCC | 690 |
| | | | |
| | | | |
| | | | |

| | | | |
|----|------|--|------|
| Db | 614 | GAGTTACTTGGGTGAAAAAAGCTTGTGGTGGGATATGAGGTCTTACATTTTCTTAT | 673 |
| Qy | 691 | TGGGACCAACACAGCTGCAGCGGAACAAAGAGG-----TGACCAGGAGAGCC | 741 |
| Db | 674 | TAGCCCAAGAGCGTGAAGGAAGCAAAATCTTGGCATTTGTCCTTACAGGATGAACACC | 733 |
| Qy | 742 | AGGCACAGGGGACATATTCCAGAGGATTTCCTAGACGTCTATTACAACTCTGACCCGTGA | 801 |
| Db | 734 | TTCTTTATGTGTGACATAATCCGACAAAGATTTTTTAGACACATATAAATAACCTGACCTTGA | 793 |
| Qy | 802 | AGACCATGATGGGCATAGAATGGGTCCATCGCTTTGTTCCTCAGCGGCGTTTGTGATGA | 861 |
| Db | 794 | AAACCATTTATGGCATTCAGGTGGTAACTGAGTTTTGCCCAATGCCAAGTACGTAAATGA | 853 |
| Qy | 862 | AAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACGTCTCTGAAGAAA | 921 |
| Db | 854 | AGACAGACATGATGTTTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAA | 913 |
| Qy | 922 | ACAGACAACCAAGTTTTTCACCTGGCTTCTGAAACTCAANTGAGTTTTCCCATCAGGCAGC | 981 |
| Db | 914 | ACCACCTCAGAGAAGTTTTTTCACAGGTTATCCCTCAATTGATAATTTTCTTATAGAGGAT | 973 |
| Qy | 982 | CATTTCAGCAAGTGGTTGTGCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCT | 1041 |
| Db | 974 | TTTACCAAAAAACCCATATTCTTACCAGAGTATCTCTTCAAGGTGTTCCTCCATCT | 1033 |
| Qy | 1042 | GCTCCGGCACCGETACGTGTTTTCTGGCGAGCTGGCAGTCAAGGTGTAACATGCTGCCA | 1101 |
| Db | 1034 | GCAGTGGGTTTGGGTTATATAAATGTCACAGATTTTGGTGCCCAAGGATCTATGAAATGATGG | 1093 |
| Qy | 1102 | AGAGCGTCCCATACATATAACTTAAACTGGAAGACGCTTTGTGGGCTCTGCCTCCAAAGGCTGA | 1161 |
| Db | 1094 | GTACGTTAAAACCCCAAGTTTGAAGATGTTATGTCTGGGATCTGTGTAATTTATTAA | 1153 |
| Qy | 1162 | ACATCAGATTTGGAGAGCTCCACTCCACGCCACCTTTTTTCCAGGGGGCTTTACGCTTCT | 1221 |
| Db | 1154 | AAGTGACATTTTCATATTCGAGAAGACACAAATCTTTTCTTCTATATAGAATCCATTGG | 1213 |
| Qy | 1222 | CCGTATGCTCTTCTCAGGAGGATCGTGGCTCGCACTTTCATCAAGCGCTCGGACTCTCTTGG | 1281 |
| Db | 1214 | ATGTCTGTCAACTGACACGTGTGATTGACGCCATGGCTTTCTTCCAAGGAGATCATCA | 1273 |
| Qy | 1282 | ACTACTGGGCAGG | 1293 |
| Db | 1274 | CTTTTGGCAGG | 1285 |

RESULT 6

US-09-990-444-208

; Sequence 208, Application US/09990444

; Publication No. US20020193300A1

; PUBLICATION NO. 0320
; GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.

APPLICANT: ASHKEWITZ, AVI J.
APPLICANT: Baker, Kevin P.

APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David

APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc

APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.

APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone

;
; APPLICANT: Ferrara, Napoleo
; APPLICANT: Fong, Sherman

APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Corritson Mary

APPLICANT: Gerritsen, Mary
APPLICANT: Goddard Andrew

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.
APPLICANT: Criminal & Christenbor

APPLICANT: Grimaldi, J. Chri

; APPLICANT: Gurney, Austin L

APPLICANT: KLJAVIN, Ivar J.

; APPLICANT: Napier, Mary A.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann

APPLICANT: Stewart, Timothy

APPLICANT: Tumas, Daniel

Db 1034 GCAGTGGGTGGTTATATAATCTCCAGAGATTGGTGCAAGGATCTATGAATGATGG 1093
QY 1102 AGAGCGTCCCATACATTAACATGGAAGACGTGTTCTGGGGCTCTGCCTCGAAAGCCTGA 1161
Db 1094 GTCACGTAACCAATCAAGTTTGAAGATGTTATGTCGGGATCTCTTTGAATTTATTA 1153
QY 1162 ACATCAGATTGGAGGAGCTCCACTCCACCGGACCTTTTTCAGGGGCTTACGCTTCT 1221
Db 1154 AAGTGAACATTATATTCAGAGACACAAATCTTTCTTCTATATAGATCCATTGG 1213
QY 1222 CCGTATGCTCTTCAGGAGATCGTGGCTGCACCTTCATCAAGCCTCGGACTCTCTGG 1281
Db 1214 ATGCTGTCAACTGAGACGTGTGATTCAGCCCATGGCTTTTCTTCCAAAGGATCATCA 1273
QY 1282 ACTACTGGCAGG 1293
Db 1274 CTTTGGCAGG 1285

RESULT 7
US-09-989-730-208
; Sequence 208, Application US/0989730
; Publication No. US20020197674A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: ROY, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730PIC69
; CURRENT APPLICATION NUMBER: US/09/989,730
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801

; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07

; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 4.7%; Score 130.4; DB 9; Length 2095;
Best Local Similarity 48.9%; Pred. No. 2.8e-25;
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;

QY 511 TTTACAAGAAACACGGGAACCTTCCTTAAGCTCCACATACAGACTCGAGCAGACACCTC 570
DB 494 TTTACAGACAAGACTTTTCACTTCACACTTCGAGAGCACTCAAACTCTCTCAAAATC 553
QY 571 CTTCTCTCGTCTGCTGTGAGCTCATCCCAACAAACAGTTGGGTGAGCGATCGCCATCC 630
DB 554 CATTTCTGCTCATCTGCTGACCTCCACCCCTTCAGATGTGAAGCCAGGCGCCATTA 613
QY 631 GCGAGAGTGGGGAAAGAGAGAGTGGTGAAGGAAAGAGCTGAAGACATCTTCTCTCC 690
DB 614 GAGTTACTTGGGGTGAAGAAAAAGTCTTGGTGGGATATGAGGTTCTTACATTTTCTTAT 673
QY 691 TGGGGACCCACAGCAGTGCAGCGGAACAAAGAGG-----TGGACAGGAGGCC 741
DB 674 TAGGCCAAGAGCTGAAAGAGGAAGACAAATCTTGCCATTGTCTTAGAGGATGAACAC 733
QY 742 AGCGACAGGGGACATATCCAGAGGATTTCCCTAGACGCTCTATTACAATCTGACCCCTGA 801
DB 734 TCTTTATGGTGACATATCCGACAGATTTTTTAGACACATATATAACCTGACCTTGA 793
QY 802 AGACCATGATGGGCATAGAATGGGTCATCGCTTTTGTCTCAGGCGGGGTTTGTGATGA 861
DB 794 AAACCATTTATGCAATTCAGGTGGGTAACTGAGCTTTTGCCCAATGCCAAGTACGTAATGA 853
QY 862 AAACGAGCTCAGACATGTTTCATCAATGTTGACTATCTGACTGAAGTCTTCTGAAGAAA 921
DB 854 AGACAGACACTGATGTTTTTCATCAATCTAGTGGCAATTTAGTGAATATCTTTTAAACCTAA 913
QY 922 ACAGAAACACAGGTTTTTTCAGTGGCTTTCTGAAACTCAATCAGTTTCCCATCAGCGAGC 981
DB 914 ACCACTCAGAGNAGTTTTTCACAGGTTATCTCTTAATGATTAATTATCTTATAGAGGAT 973
QY 982 CATTACAGCAAGTGGTTTGTCTAGTAATCTGAATATCCGTGGGACAGGTACCCACCATTTCT 1041
DB 974 TTTACCAAAAAACCCATATTTCTTACCAGGAGTATCTTTCAGAGGTGTTCCCTCCATCT 1033
QY 1042 GCTCCGGCAGCGGTACGTGTTTCTGGGAGGTGGCGAGTCAAGTGTACAATGTCTCCA 1101
DB 1034 GCAGTGGGTGGGTATATAATGTCCAGAGATTTGGTCCAGAGATCTATGAATGATGG 1093
QY 1102 AGAGCGTCCCATACATTAACCTGGAAGAGCTGTTTGTGGGCTCTGCCCTCGAAGGCTCA 1161
DB 1094 GTCAGTAAACCCCATCAAGTTTGAAGATGTTTATGTCGGGATCTGTTTGAATTTATTA 1153
QY 1162 ACATCAGATTGGAGAGCTCCACTCCAGCCGACCTTTTTTCCAGGGGCTTTACGCTTCT 1221
DB 1154 AAGTGAACATTCATATTCAGAGAAGACACAAATCTTTTCTTCTATATAGAATCCATTGG 1213
QY 1222 CCGTATGCTCTTCAGGAGGATCGTGGCTGCCACTTCATCAAGCCTCGGACTCTCTTGG 1281
DB 1214 ATGCTGTCAACTGAGAGCTGTGATTGCAGGCCCATGGCTTTTCTTCCAAGGAGATCATCA 1273
QY 1282 ACTACTGGCAGG 1293
DB 1274 CTTTTTGGCAGG 1285

RESULT 8

US-09-990-436-208
; Sequence 208, Application US/09990436
; Publication No. US20020198148A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730PIC14
CURRENT APPLICATION NUMBER: US/09/990,436
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444

PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 4.7%; Score 130.4; DB 9; Length 2095;
Best Local Similarity 48.9%; Pred. No. 2.8e-25;
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;

QY 511 TTTACAGAAGACGGGAAGTCTTAAAGTCCCGATACAGACTGCGAGCGACAGACCTC 570
DB 494 TTTACAGAAGACCTTCACTCACACTTCGAGAGCAATTCAAACTGCTCTCATCAAAATC 553
QY 571 CTTCTCGTCGTCGTTGGTACCTCTCCACAAAGATGGCTGAGCGCATGCCATCC 630
DB 554 CATTTCTGTCATCTTGGTGACCTCCACCTTCAGATGTGAAGCCAGCGGCGCATTA 613
QY 631 GCGAGACGTGGGGAAGAGAGGATGGTGAAGGAAGCAGCTCAAGACATCTTCTCTCC 690
DB 614 GAGTTACTTGGGTGAAAAAAGTCTTGGTGGGATATGAGTCTTACATTTTCTTAT 673
QY 691 TGGGGACCAACCACTGCGCGGGAACGAGG-----TGGACCAAGAGGCC 741
DB 674 TAGGCCAAGAGGCTGAAAGAGGAAGCAAAATGTGGCATTTGCTTAGAGGATGAACACC 733
QY 742 AGCGACACGGGACATTTATCCGAAGGATTTCTAGACGTCTATTACATCTGACCTGA 801
DB 734 TTCTTTATGGTGACATAATCCGACAGAGATTTTACACACATATATACCTGACCTTGA 793
QY 802 AGACCATGATGGGCATAGAAATGGGTCCATCGCTTTTGTCTCTCAGCGGCGGTGTGATGA 861

DB 794 AAACCATTTATGGCATTCAGGTGGGTAACTGAGTTTGGCCCAATCCCAATGACGTAAAGA 853
QY 862 AAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAAGTGTCTTGAAGAAA 921
DB 854 AGACAGACACTGATGTTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAA 913
QY 922 ACAGAACACCAAGGTTTTTTCACCTGCTTCTTGAAGAACTCAATGAGTTTCCCATCAGGCAGC 981
DB 914 ACCACTCAGAGAAGTTTTTTCACAGGTTATCCTCTAAATTGATAAATTTATCTATAGAGGAT 973
QY 982 CATTACAGCAAGTGGTTTGTCTAGTAAATCTGAATATCCGTGGCAGACAGTACCCACCATCT 1041
DB 974 TTTACCAAAAACCCATATTTCTTACCAGAGATATCTTTCAAGGTGTTCCCTCCATCT 1033
QY 1042 GTCGCGCACCGGCTACGTGTTTCTGCGGACGTGGCGAGTCAGGTGACAAATGCTCTCCA 1101
DB 1034 GCAGTGGTGGTTATATAATGTCACAGATTTGGTCCCAAGGATCTATGAATGATGG 1093
QY 1102 AGACGCTCCCATACATTTAACTGGAAGACGTGTTTGTGGGCTCTGCGCTCGAAGGCTGA 1161
DB 1094 GTCACGTAAACCCATCAAGTTTGAAGATGTTTATGTCGGATCTGTTGAATTTATTAA 1153
QY 1162 ACATCAGATTGGAGAGCTCCACTCCAGCCGACCTTTTCCAGGGGCTTACGCTTCT 1221
DB 1154 AAGTGAACATTCATATTCAGAGACACAAATCTTTCTTATATAGATTCATTTGG 1213
QY 1222 CCGTATGCTCTTCAGGAGGATCGTGGCTGCTGCTCATCAAGCTCGGACTCTTTGG 1281
DB 1214 ATGCTGTCAACTGAGAGCTGTGATTCAGGCCATGGCTTTCTTCCAGGAGATCA 1273
QY 1282 ACTACTGGCAGG 1293
DB 1274 CTTTGGCAGG 1285

RESULT 9

US-09-991-181-208
; Sequence 208, Application US/09991181
; Publication No. US20020197615A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC53
; CURRENT APPLICATION NUMBER: US/09/991,181
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250

| | |
|----|-------------------------------------|
| 1 | PRIOR APPLICATION NUMBER: 60/089512 |
| 2 | PRIOR FILING DATE: 1998-06-16 |
| 3 | PRIOR APPLICATION NUMBER: 60/089514 |
| 4 | PRIOR FILING DATE: 1998-06-16 |
| 5 | PRIOR APPLICATION NUMBER: 60/089532 |
| 6 | PRIOR FILING DATE: 1998-06-17 |
| 7 | PRIOR APPLICATION NUMBER: 60/089538 |
| 8 | PRIOR FILING DATE: 1998-06-17 |
| 9 | PRIOR APPLICATION NUMBER: 60/089598 |
| 10 | PRIOR FILING DATE: 1998-06-17 |
| 11 | PRIOR APPLICATION NUMBER: 60/089599 |
| 12 | PRIOR FILING DATE: 1998-06-17 |
| 13 | PRIOR APPLICATION NUMBER: 60/089600 |
| 14 | PRIOR FILING DATE: 1998-06-17 |
| 15 | PRIOR APPLICATION NUMBER: 60/089653 |
| 16 | PRIOR FILING DATE: 1998-06-17 |
| 17 | PRIOR APPLICATION NUMBER: 60/089801 |
| 18 | PRIOR FILING DATE: 1998-06-18 |
| 19 | PRIOR APPLICATION NUMBER: 60/089907 |
| 20 | PRIOR FILING DATE: 1998-06-18 |
| 21 | PRIOR APPLICATION NUMBER: 60/089908 |
| 22 | PRIOR FILING DATE: 1998-06-18 |
| 23 | PRIOR APPLICATION NUMBER: 60/089947 |
| 24 | PRIOR FILING DATE: 1998-06-19 |
| 25 | PRIOR APPLICATION NUMBER: 60/090246 |
| 26 | PRIOR FILING DATE: 1998-06-22 |
| 27 | PRIOR APPLICATION NUMBER: 60/090252 |
| 28 | PRIOR FILING DATE: 1998-06-22 |
| 29 | PRIOR APPLICATION NUMBER: 60/090254 |
| 30 | PRIOR FILING DATE: 1998-06-22 |
| 31 | PRIOR APPLICATION NUMBER: 60/090349 |
| 32 | PRIOR FILING DATE: 1998-06-23 |
| 33 | PRIOR APPLICATION NUMBER: 60/090355 |
| 34 | PRIOR FILING DATE: 1998-06-23 |
| 35 | PRIOR APPLICATION NUMBER: 60/090429 |
| 36 | PRIOR FILING DATE: 1998-06-24 |
| 37 | PRIOR APPLICATION NUMBER: 60/090431 |
| 38 | PRIOR FILING DATE: 1998-06-24 |
| 39 | PRIOR APPLICATION NUMBER: 60/090435 |
| 40 | PRIOR FILING DATE: 1998-06-24 |
| 41 | PRIOR APPLICATION NUMBER: 60/090444 |
| 42 | PRIOR FILING DATE: 1998-06-24 |
| 43 | PRIOR APPLICATION NUMBER: 60/090445 |
| 44 | PRIOR FILING DATE: 1998-06-24 |
| 45 | PRIOR APPLICATION NUMBER: 60/090472 |
| 46 | PRIOR FILING DATE: 1998-06-24 |
| 47 | PRIOR APPLICATION NUMBER: 60/090535 |
| 48 | PRIOR FILING DATE: 1998-06-24 |
| 49 | PRIOR APPLICATION NUMBER: 60/090540 |
| 50 | PRIOR FILING DATE: 1998-06-24 |
| 51 | PRIOR APPLICATION NUMBER: 60/090542 |
| 52 | PRIOR FILING DATE: 1998-06-24 |
| 53 | PRIOR APPLICATION NUMBER: 60/090557 |
| 54 | PRIOR FILING DATE: 1998-06-24 |
| 55 | PRIOR APPLICATION NUMBER: 60/090676 |
| 56 | PRIOR FILING DATE: 1998-06-25 |
| 57 | PRIOR APPLICATION NUMBER: 60/090694 |
| 58 | PRIOR FILING DATE: 1998-06-25 |
| 59 | PRIOR APPLICATION NUMBER: 60/090695 |
| 60 | PRIOR FILING DATE: 1998-06-25 |
| 61 | PRIOR APPLICATION NUMBER: 60/090696 |
| 62 | PRIOR FILING DATE: 1998-06-25 |
| 63 | PRIOR APPLICATION NUMBER: 60/090862 |
| 64 | PRIOR FILING DATE: 1998-06-26 |
| 65 | PRIOR APPLICATION NUMBER: 60/090863 |

; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24

; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 4.7%: Score 130.4; DB 9; Length 2095;
Best Local Similarity 48.9%: Pred. No. 2.8e-25;
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;

QY 511 TTTACAAGAAAGCGGGAACCTTCCTTAAGCTCCAGATACAGACTCGAGGCGAGACCTTC 570
DB 494 TTTACAGACAAGACTTTCACACTTCGAGAGCATTCAAACTGCTCTCATCAAAATC 553
QY 571 CCTCTCTGCTGCTGCTGACCTCATCCACAAACAGTTGGCTGAGCGCATGGCCATCC 630
DB 554 CATTTCTGCTGCTGCTGACCTCCACCTTCAGATGTGAAGCCAGGCGGCGCTTA 613
QY 631 GGCAGACGTGGGGAAAGAGAGGATGGTGAAGGAAAGCAGCTGAAGACATCTTCTCTCC 690
DB 614 GAGTTACTTGGGGTCAAAAAAAGTCTTGGTGGGATATGAGGTTCTTACATTTTCTTAT 673
QY 691 TGGGACCCAGCAGCTGCGCGGGAACGAAGAGG-----TGGACAGGAGGCC 741
DB 674 TAGGCCAAGAGGCTGAAAAGGAAGACAAATGTTGGCATTTGCTTAGAGGATGAACACC 733
QY 742 AGCGACACGGGACATTTATCCAGAAGGATTTCTTAGAGCTCTATTACATCTGACCTCA 801
DB 734 TCTTTATGCTGACATATATCCGCAAGATTTTTTAGACACATATATAACCTGACCTTGA 793
QY 802 AGACCATGATGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGCGGGGTTTGTGATGA 861
DB 794 AAACCATTTAGCATTCAGGTGGGTAACTGAGTTTTGGCCCCCAATGCCAGTACGTAAATGA 853
QY 862 AAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAAGTCTTCTTGAAGAAA 921
DB 854 AGACAGACACTGATGTTTTTCATCACTACTGGCAATTTAGTGAAGTATCTTTTAAACCTAA 913
QY 922 ACAGAAACCAACGAGTTTTTTCACCTGCTTCTTGAACACTCAATGAGTTTCCCATCAGCGAGC 981
DB 914 ACCACTCAGAGAAGTTTTTTTTCACAGGTTATCTCTAAATTTGATAAATTTATTTCTTATAGAGAT 973

; Publication No. US20030008352A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430RIC42

; CURRENT APPLICATION NUMBER: US/10/174,590

; CURRENT FILING DATE: 2002-06-18

; Prior application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 161

; LENGTH: 2095

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-174-590-161

Query Match 4.7%; Score 130.4; DB 9; Length 2095;

Best Local Similarity 48.9%; Pred. No. 2.8e-25;

Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;

Qy 511 TTTACAAGAAAGACGGGAACCTTCTTAAGCTCCACAGATACAGACTGCAGGACAGACCTC 570

Db 494 TTTACAGACAAGACTTTTCACTTCACACTTCGAGAGCATTCACAACTGCTCTCATCAAAATC 553

Qy 571 CTTCTCCTCGTCTGCTGTGACCTCATCCACAAACAGTTGGCTGAGCGCATGGCCATCC 630

Db 554 CATTTCTGTCTATCTGTGTGACCTCCACCTTCAGATGTGAAGCCAGGCGCATTA 613

Qy 631 GGCAGACTGGGGGAAGAGAGGATGGTGAAGGGAACAGCTGAAGACATCTTCTCTCC 690

Db 614 GAGTTACTTGGGTGAAAAAAGTCTTGGTGGGATATGAGTTCTTACATTTTCTTAT 673

Qy 691 TGGGACCACCCAGCAGTGCAGCGGAACGAAGAGG-----TGGACCAGGAGAGCC 741

Db 674 TAGGCCAAGAGGCTGAAAGAGGACAAAATGTTGGCATTTGCTTAGAGGATGAACACC 733

Qy 742 AGCGACAGCGGGACATATCCAGAGGATTTCTTAGAGCTCTATACAACTGACCCCTGA 801

Db 734 TTCTTTATGGTGACATAATCCGACAAGATTTTGTAGACACATATAAATACCTGACCTGA 793

Qy 802 AGACCATGATGGCATAGAATGGGTGCCATCGCTTTGTGCTCAGCGCGGCTTGTGATGA 861

Db 794 AAACCATATGGCATTGAGTGGGTAACTGAGTTTGGCCCAATGCCAAGTACGTAATGA 853

Qy 862 AAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTCAACTGCTTCTGAAGAAA 921

Db 854 AGACAGACACTGATGTTTTTCACTCAATACTGGCAATTTAGTGAAGTATCTTTAAACCTAA 913

Qy 922 ACAGAACACACAGTTTTTCACTGCTTCTTGAACACTCAATGAGTTTCCCATCGGAGC 981

Db 914 ACCACTCAGAGAGTTTTTTCACAGGTTATCCTCTAATTTGATAATTTCTTATAGAGAT 973

Qy 982 CATTCAGCAAGTGGTTGTCTAGTAAATCTGAATATCCCTGGGACAGGTACCCACCATCT 1041

Db 974 TTTACCAAAAACCCATATTTCTTACCAGGAGTATCCTTTCAAGGTGTCCCTCCATACT 1033

Qy 1042 GCTCCGACCGGTACGTGTTTTCTGGCGACGTGGGAGTCAGGTGTACAAATGTCCTCA 1101

Db 1034 GCATGGGTTGGGTTATATATATGTCACAGATTTGGTGCCCAAGGTCTATGAAATGATGG 1093

Qy 1102 AGACGGTCCCATACATTAACCTGGAAGACGTGTTGTGGGGCTCTGCTCCAAAGGCTGA 1161

Db 1094 GTCACGTAACCCATCAAGTTTGAAGATGTTTATGTCGGGATCTGTTTGAATTTATTA 1153

Qy 1162 ACATCAGATGGAGGAGCTCCACTCCACGCCGACCTTTTTTCCAGGGGCTTACGCTTCT 1221

Db 1154 AAGTGAACATTATATTTCCAGAGACACAAATCTTTTCTCTATATAGAAATCCATTGG 1213

Qy 1222 CCGTATGCTCTTTCAGGAGGATCGTGGCTGACACTTCATCAAGCCCTCGGACTCTCTTGG 1281

Db 1214 ATGCTGTCAACTGAGACGCTGTGATGACGCCCATGGCTTTCTTCCAAGGAGATCATCA 1273

Qy 1282 ACTACTGGCAGG 1293

Db 1274 CTTTTTGGCAGG 1285

RESULT 14

US-10-176-758-161

; Sequence 161, Application US/10176758

; Publication No. US20030008353A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430RIC104

; CURRENT APPLICATION NUMBER: US/10/176,758

; CURRENT FILING DATE: 2002-06-21

; Prior application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 161

; LENGTH: 2095

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-176-758-161

Query Match 4.7%; Score 130.4; DB 9; Length 2095;

Best Local Similarity 48.9%; Pred. No. 2.8e-25;

Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;

Qy 511 TTTACAAGAAAGACGGGAACCTTCTTAAGCTCCACAGATACAGACTGCAGGACAGACCTC 570

Db 494 TTTACAGACAAGACTTTTCACTTCACACTTCGAGAGCATTCACAACTGCTCTCATCAAAATC 553

Qy 571 CTTCTCCTCGTCTGCTGTGACCTCATCCACAAACAGTTGGCTGAGCGCATGGCCATCC 630

Db 554 CATTTCTGTCTATCTGTGTGACCTCCACCTTCAGATGTGAAGCCAGGCGCATTA 613

Qy 631 GGCAGACTGGGGGAAGAGAGGATGGTGAAGGGAAGAGCTGAAGACATCTTCTCTCC 690

Db 614 GAGTTACTTGGGTGAAAAAAGTCTTGGTGGGATATGAGGTTCTTACATTTTCTTAT 673

Qy 691 TGGGACCACCCAGCAGTGCAGCGGAACGAAGAGG-----TGGACCAGGAGAGCC 741

Db 674 TAGGCCAAGAGGCTGAAAGAGGACAAAATCTTGGCATTTGCTTAGAGGATGAACACC 733

Qy 742 AGCGACAGCGGGACATATTCAGAGGATTTCTTAGACCTCTATTACAACTGACCCCTGA 801

Db 734 TTCTTTATGGTGACATAATCCGACAAGATTTTGTAGACACATATAAATACCTGACCTGA 793

Qy 802 AGACCATGATGGCATAGAATGGGTCCATCGCTTTGTCTCAGCGGGCTTGTGATGA 861

Db 794 AAACCATATGGCATTGAGTGGGTAACTGAGTTTGGCCCAATGCCAAGTACGTAATGA 853

Qy 862 AAACAGACTCAGACATGTTTCAATCAATGTTGACTATCTGACTCAACTGCTTCTGAAGAAA 921

Db 854 AGACAGACACTGATGTTTTCATCAATACTGGCAATTAGTGAAGATATCTTTTAAACCTAA 913
QY 922 ACAGACAACACAGGTTTTCACCTGGCTCTTGAACACTCAATGAGTTTCCCATCAGGCAGC 981
Db 914 ACCACTCAGAGAAGTTTTCACAGGTTATCCTCTAATGATAATATATCTATAGAGGAT 973
QY 982 CATTCAGCAAGTGGTTTGTTCAGTAATCTGAATATCCGTGGGACAGGTACCCACCATCT 1041
Db 974 TTTACCAAAAACCCATATTTCTTACCAGGAGTATCCTTCAAGGTGTTCCCTCCATCT 1033
QY 1042 GCTCCGGCACCCTAGCTGTTTCTGCGCAGCTGGGAGTCAAGTGTACAATGTCTCCA 1101
Db 1034 GCAGTGGGTGGGTATATATATGTCCAGAGATTGGTGCCCAAGGATCTATGAATGATGG 1093
QY 1102 AGAGCTCCCATACATATAACTGGAAGACGCTGTTGTGGGGCTCTGCTCGAAAGGCTGA 1161
Db 1094 GTCACGTAAAACCCATCAAGTTGAAGATCTTATGTGCGGATCTGTTGAATTTATTA 1153
QY 1162 ACATCAGATTGGAGAGCTCCACTCCAGCCGACCTTTTTCAGGGGGCTTACGCTTCT 1221
Db 1154 AAGTGAACATTCATATCCAGAAGACACAAATCTTTTCTTATATAGAATCCATTGG 1213
QY 1222 CCCTATGCTCTTCAGGAGATCGTGGCTGGCCCTGCCACTTCATCAAGCTCGGACTCTCTGG 1281
Db 1214 ATGTCTGTCAACTGAGACGTGTGATTGACGCCCATGGCTTTTCTTCAAGGAGATCATCA 1273
QY 1282 ACTACTGGCAGG 1293
Db 1274 CTTTTTGGCAGG 1285

RESULT 15
US-10-063-616-35
; Sequence 35, Application US/10063616
; Publication No. US20030013855A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,616
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 35
; LENGTH: 2095
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-616-35

Query Match 4.7%; Score 130.4; DB 9; Length 2095;
Best Local Similarity 48.9%; Pred. No. 2.8e-25;
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;
QY 511 TTTACAGAAAGACGGGAACCTTCTTAAGCTCCAGATACAGACTGCAGGACAGACACCTC 570
Db 494 TTTACAGACAAGACTTTCACTTCACACTTCGAGAGCATTCAAACTGCTCTCATCAAAATC 553
QY 571 CCTTCCTGCTGCTGGTGACCTCATCCACAAAGTGTGGCTGAGCGCATGGCCATCC 630
Db 554 CATTTCTGTGTCATCTGGTGACCTCCACCTTCAGATGTCGAAAGCCAGGAGCCCATTA 613
QY 631 GGCAGACGTGGGGGAAAGAGAGAGTGGTGAAGGAAAGACGTGAAGACATCTCTCTCC 690
Db 614 GAGTTACTTGGGGTGAAGAAAGAGTCTGGTGGGATATGAGTCTTACATTTTCTTAT 673

QY 691 TGGGACCACACAGCAGTGCAGCGAAACGAAAGAGG-----TGGACAGGAGAGACC 741
Db 674 TAGCCCAAGAGGCTGAAAAGGAAGACAAATGTTGGCAATTCTCTTAGAGGATGAACACC 733
QY 742 AGCAGACAGGGGACATTTATCCAGAAAGGATTTCCTAGACGTCTATTACAAATCTGACCCCTGA 801
Db 734 TTTCTTTATGGTGACATAATCCGACAAGATTTTTTAGACACATATAAATAACCTGACCTTGA 793
QY 802 AGACCATGATGGGATAGAAATGGGTCCATCGCTTTTGTCTCCAGCGCGCTTTGTGATGA 861
Db 794 AAACCATTTATGGCAATTCAGGTGGGTAATCGAGTTTGGCCCAATGCCAAGTACGTAATGA 853
QY 862 AAACAGACTCAGACATGTTCAATCAATGTTGACTATCTGACTGAACTGCTCTGGAAGAAA 921
Db 854 AGACAGACACTGATGTTTTCATCAATCTGCAATTAAGTCAAGTATCTTTTAAACCTAA 913
QY 922 ACAGAAACAACAGGTTTTTTCACGTGCTTTTGAACCTCAATGAGTTTCCCATCAGGCAGC 981
Db 914 ACCACTCAGAGAAGTTTTTTCACAGGTTATCCTCTAATTGATAAATATTTCCTATAGAGGAT 973
QY 982 CATTCAGCAAGTGGTTTGTCTAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATCT 1041
Db 974 TTTACCAAAAACCCATATTTTACCAGGAGTATCCTTTTCAAGGTGTTCCCTCCATCT 1033
QY 1042 GCTCCGGCACCCTGCTACGTGTTTCTGCGACGTGGGAGTCAAGTGTACAAATGTCTCCA 1101
Db 1034 GCAGTGGGTGGGTATATATATGTCAGAGATTTGGTGCCCAAGGATCTATGAAATGATGG 1093
QY 1102 AGAGCGTCCCATACATTTAAACTGGAAGACGCTGTTTGTGGGGCTCTGCTCGAAAGGCTGA 1161
Db 1094 GTCACGTAAAACCCATCAAGTTTCAAGATGTTTATGTCGGGATCTGTTTGAATTTATTA 1153
QY 1162 ACATCAGATTGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCT 1221
Db 1154 AAGTGAACATTCATATTTCCAGAAGACACAAATCTTTTCTTATATAGAATCCATTGG 1213
QY 1222 CCGTATGCTCTTTCAGGAGGATCGTGGCTGCCACTTCATCAAGCTCGGACTCTCTTGG 1281
Db 1214 ATGTCTGTCAACTGAGACGTGTGATTGCGCCCATGGCTTTTCTTCCAAGGAGATCATCA 1273
QY 1282 ACTACTGGCAGG 1293
Db 1274 CTTTTTGGCAGG 1285

Search completed: April 12, 2003, 03:33:18
Job time : 184.543 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2003, 12:08:21 ; Search time 33 Seconds

(without alignments)

1935.594 Million cell updates/sec

Title: US-09-914-152-1

Perfect score: 1657

Sequence: 1 MAPPKRLMVICLLVLGALC.....TLDDYWALENSRGEDCPVP 310

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1 | 1541.5 | 93.0 | 302 | 6 | Q9N292 |
| 2 | 589 | 35.5 | 326 | 11 | Q54904 |
| 3 | 587 | 35.4 | 326 | 4 | Q9Y526 |
| 4 | 587 | 35.4 | 326 | 6 | Q9MYM7 |
| 5 | 576 | 34.8 | 305 | 11 | Q920V5 |
| 6 | 576 | 34.8 | 305 | 11 | Q91V52 |
| 7 | 569 | 34.3 | 422 | 4 | Q43825 |
| 8 | 567 | 34.2 | 409 | 11 | Q920V4 |
| 9 | 567 | 34.2 | 409 | 11 | Q920V3 |
| 10 | 567 | 34.2 | 409 | 11 | Q920V2 |
| 11 | 567 | 34.2 | 409 | 11 | Q91VE9 |
| 12 | 567 | 34.2 | 409 | 11 | Q91V58 |
| 13 | 567 | 34.2 | 409 | 11 | Q91V19 |
| 14 | 567 | 34.2 | 422 | 11 | Q54905 |
| 15 | 565 | 34.1 | 422 | 4 | Q9B2Q9 |
| 16 | 558 | 33.7 | 331 | 11 | Q920V1 |

| | | | | | | |
|----|-------|------|-----|----|--------|---------------------|
| 17 | 558 | 33.7 | 331 | 11 | Q920V0 | Q920V0 mus musculus |
| 18 | 558 | 33.7 | 331 | 11 | Q91V72 | Q91V72 mus musculus |
| 19 | 558 | 33.7 | 331 | 11 | O54906 | O54906 mus spicile |
| 20 | 555.5 | 33.5 | 331 | 4 | O75752 | O75752 homo sapien |
| 21 | 529.5 | 32.0 | 255 | 11 | O9CTE5 | O9CTE5 mus musculus |
| 22 | 434.5 | 26.2 | 412 | 13 | O8UWM2 | O8UWM2 brachydanio |
| 23 | 429 | 25.9 | 204 | 4 | O8TDY1 | O8TDY1 homo sapien |
| 24 | 416 | 25.1 | 372 | 4 | O9C0J2 | O9C0J2 homo sapien |
| 25 | 407.5 | 24.6 | 390 | 13 | O8UWM1 | O8UWM1 brachydanio |
| 26 | 403.5 | 24.4 | 372 | 11 | Q9D722 | Q9D722 mus musculus |
| 27 | 402.5 | 24.3 | 372 | 11 | O8R0U2 | O8R0U2 mus musculus |
| 28 | 396 | 23.9 | 378 | 4 | O9BYG0 | O9BYG0 homo sapien |
| 29 | 393 | 23.7 | 371 | 11 | O88178 | O88178 ratus norv |
| 30 | 388 | 23.4 | 370 | 11 | Q920U7 | Q920U7 mus spicile |
| 31 | 388 | 23.4 | 370 | 11 | Q91VC1 | Q91VC1 mus musculus |
| 32 | 388 | 23.4 | 371 | 11 | Q920F0 | Q920F0 mus musculus |
| 33 | 388 | 23.4 | 378 | 4 | O96024 | O96024 homo sapien |
| 34 | 388 | 23.4 | 406 | 13 | O8UWM0 | O8UWM0 brachydanio |
| 35 | 386 | 23.3 | 418 | 13 | O8UWM4 | O8UWM4 brachydanio |
| 36 | 383.5 | 23.1 | 384 | 4 | O8TDX1 | O8TDX1 homo sapien |
| 37 | 383 | 23.1 | 370 | 11 | Q920U8 | Q920U8 mus musculus |
| 38 | 382.5 | 23.0 | 277 | 4 | O8TAZ4 | O8TAZ4 homo sapien |
| 39 | 381.5 | 23.0 | 574 | 5 | O9VLS8 | O9VLS8 drosophila |
| 40 | 381.5 | 23.0 | 585 | 5 | O9STQ8 | O9STQ8 drosophila |
| 41 | 380 | 22.9 | 370 | 11 | Q920U9 | Q920U9 mus musculus |
| 42 | 364 | 22.0 | 412 | 13 | O8UWM3 | O8UWM3 brachydanio |
| 43 | 362 | 21.8 | 399 | 11 | O8V116 | O8V116 mus musculus |
| 44 | 359.5 | 21.7 | 377 | 4 | O96EK0 | O96EK0 homo sapien |
| 45 | 359 | 21.7 | 397 | 11 | Q91V18 | Q91V18 mus musculus |

ALIGNMENTS

RESULT 1

Q9N292 ID Q9N292 PRELIMINARY: PRT; 302 AA.

AC Q9N292; DT 01-OCT-2000 (TREMBLrel. 15, Created) DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update) DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update) DE UDP-Gal:GlcNAc beta1,3-galactosyltransferase 5 (fragment). GN BETAL1,3-GALT 5.

OS Pongo pygmaeus (Orangutan). OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pongo. OX NCBI_TaxID=9600; RN [1]

RP SEQUENCE FROM N.A. RC STRAIN-ORAN-POL7; RA Liu Y., Saitou N.; RT "Silver Project";

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases. DR EMBL; AB041417; BAA94502.1; --

DR InterPro; IPR002659; Galactosyl_T. DR Pfam; PF01762; Galactosyl_T; 1.

KW Glycosyltransferase; Transferase. FT NON_TER 302 302

SQ SEQUENCE 302 AA; 51A555C864465A90 CRC64;

Query Match 93.0%; Score 1541.5; DB 6; Length 302; Best Local Similarity 96.4%; Pred. No. 1.8e-137; Matches 291; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MAPPKRLMVICLLVLGALCLYFSMYSLNPFREQSFVYKK-DGNFLKLPDTCRQTPPFL 59
Db 1 MAPPKRLIICLLVLGALCLYFSMYSLNPFREQSFVYKKEDGNFLKLPDTCRQTPPFL 60

QY 60 VLLVTSKSHKQLAERMAIRTWGKRMVKQKLTFFLLGTSSAAETKEDVESORHGDI 119
Db 61 VLLVTSKSHKQLAERMAIRTWGKRMVKQKLTFFLLGTSSAAETKEDVESORHGDI 120

QY 120 IQKDFLDVYNYNLTLMGMGIEWHRCFQPAAYFNKTDSDNFINVDTLLELLKKNTRTF 179

```

Db 121 IOKDFLOWYVNTLTKTMGMSEWHRFCQAAAFVMTKDSMDNFINDVYITELLKKNRTFR 180
Qy 180 FTGFLKNEPIROPFSKWFYSKSEYPMWDRYPPFCSGTGYVFGSDVQVNVSKSVPI 239
Db 181 FTGFLKNEPIROPFSKWFYSKSEYPMWDRYPPFCSGTGYVFGSDVQVNVSKSVPI 240
Qy 240 KLEDVFGVGLCLERLNIRLEELHSOPTFPFGGLRSVCLFRIVACHFKPRTLDDYQAL 299
Db 241 KLEDVFGVGLCLERLNIRLEELHSOPTFPFGGLRSVCLFRIVACHFKPRTLDDYQAL 300
Qy 300 EN 301
Db 301 EN 302

RESULT 2
O54904 PRELIMINARY; PRT; 326 AA.
AC O54904
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE UDP-galactose:BETA-N-acetyl-glucosamine-BETA-1,3-galactosyltransferase
DE II (EC 2.4.1.-) (UDP-GAL:BETAGLGNAC BETA 1,3-GALACTOSYLTRANSFERASE-I)
DE (BETA3GALT-I).
GN B3GTL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN 1
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RX STRAIN=129/SVJ;
RX MEDLINE=98079027; PubMed=9417047;
RA Hennen T., Dinter A., Kuhnert P., Mattu T.S., Rudd P.M., Berger E.G.;
RT "Genomic cloning and expression of three murine UDP-galactose: beta-N-
acetylglucosamine beta1,3-galactosyltransferase genes.";
RL J. Biol. Chem. 273:58-65(1998).
CC -1- FUNCTION: THIS PROTEIN IS RESPONSIBLE FOR THE SYNTHESIS OF
CC COMPLEX-TYPE N-LINKED OLIGOSACCHARIDES IN MANY GLYCOPROTEINS A AS
CC WELL AS THE CARBOHYDRATE MOETIES OF GLYCOLIPIDS. GALT-I, GALT-II
CC AND GALT-III ENZYMES ACT ON THE SAME ACCEPTOR SUBSTRATES BUT WITH
CC DIFFERENT AFFINITY.
CC -1- CATALYTIC ACTIVITY: UDP-GALACTOSE + N-ACETYL-BETA-D-GLUCOSAMINYL-
CC GLYCOPETIDE -> UDP + BETA-D-GALACTOSYL-1,3-N-ACETYL-BETA-D-
CC GLUCOSAMINYLGLYCOPETIDE.
CC -1- COFACTOR: MANGANESE.
CC -1- PATHWAY: GLYCOSYLATION.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND TO A LESSER EXTENT IN
CC HEART, KIDNEY, LIVER, LUNG, OVARY, TESTIS, UTERUS, COLON, STOMACH,
CC BONE MARROW, THYMUS, SPLEEN AND LYMPH NODES.
DR EMBL: AF029790; AAC53523.1; -
DR MGD; MGI:1349403; B3galt1.
DR InterPro: IPR002659; Galactosyl_T.
DR Pfam: PF01762; Galactosyl_T; 1.
KW Transferase; Glycosyltransferase; Transmembrane; Signal-anchor;
KW Glycoprotein; Golgi stack.
FT DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 7 25 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DOMAIN 26 326 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 326 AA; 37979 MW; 9825D6869EFA3D CRC64;

Query Match 35.5%; Score 589; DB 11; Length 326;
Best Local Similarity 38.6%; Pred. No. 1.9e-47;
Matches 127; Conservative 52; Mismatches 108; Indels 42; Gaps 7;

Qy 12 CLAVIGALCALFYMSYLSN-----PFKEQSVYKK--DGN-----FL 45
Db 7 CLYLVSWCVASALWYLSITRPTSSVTGSKPFSHLTVARKNFTFGNIRTPINPHSPEFL 66

```

```

Qy 46 KLPTDCRCQTPPFLVLLTSSHKQLAERMAIROTWGERKRWGKOLKTFLLGTTSSAAE 105
Db 67 INEPNCKEKNIPFVLILSTHKEFDARQAIRETWGDENNFGIKIATIFLLGKNADPVL 126
Qy 106 TKEVDQESQRHGDITOKOFLDYYVNTLTKTMGMSEWHRFCQAAAFVMTKDSMDNFINDV 165
Db 127 NOMVQESQIFHDIIVEDFIDSYHNLTLLTKLMGRVATFCSKAKYVMTKDSIDFYNDN 186
Qy 166 LTELKLNKR--TTRFFTGFLKNEPIROPFSKWFYSKSEYPMWDRYPPFCSGTGYVFG 223
Db 187 LIYKLFSTPRRRIYFTGYV-INGGPIRVKSRMPRLYDPSNYPFCSGTGYIFA 245
Qy 224 DVASQVNVSKSVPIKLEDFVGLCLERLNIRLEELHSOPTFPFGG-----LRFSVCLF 278
Db 246 DVAELIYKTLTSLTRLLHLEDVYVGLCKRKGTH-----PFQNSGFNHNKMYSLCRY 297
Qy 279 RRVACHFIKPRTLDDYQALNSRGDC 307
Db 298 RRVITVHQISPEEMHRIWNDSKKHLRC 326

RESULT 3
Q9Y5Z6 PRELIMINARY; PRT; 326 AA.
AC Q9Y5Z6;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE UDP-galactose:BETA-N-acetyl-glucosamine-BETA-1,3-galactosyltransferase
DE I (EC 2.4.1.-) (BETA3GALT1).
GN B3GALT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX TISSUE-COLON ADENOCARCINOMA;
RX MEDLINE=99283878; PubMed=10356986;
RA Bardoní A., Valli M., Trinchera M.;
RT "Differential expression of beta1,3galactosyltransferases in human
RT colon cells derived from adenocarcinomas or normal mucosa.";
RL FEBS Lett. 451:75-80(1999).
RN 2
FUNCTION, AND TISSUE SPECIFICITY.
RX MEDLINE=98250717; PubMed=9582303;
RA Amado M., Almeida R., Carneiro F., Levery S.B., Holmes E.H.,
RA Nomoto M., Hollingsworth M.A., Hassan H., Schwientek T., Nielsen P.A.,
RA Bennett E.P., Clausen H.;
RT "A family of human beta3-galactosyltransferases. Characterization of
RT four members of a UDP-galactose:beta-N-acetyl-glucosamine/beta-
RT nacyl-galactosamine beta-1,3-galactosyltransferase family.";
RL J. Biol. Chem. 273:12770-12778(1998).
RN 3
SEQUENCE FROM N.A.
RA Liu Y., Saitou N.;
RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS PROTEIN IS RESPONSIBLE FOR THE SYNTHESIS OF
CC COMPLEX-TYPE N-LINKED OLIGOSACCHARIDES IN MANY GLYCOPROTEINS A AS
CC WELL AS THE CARBOHYDRATE MOETIES OF GLYCOLIPIDS. GALT1 AND GALT2
CC ENZYMES ACT ON THE SAME ACCEPTOR SUBSTRATES BUT WITH DIFFERENT
CC AFFINITY.
CC -1- CATALYTIC ACTIVITY: UDP-GALACTOSE + N-ACETYL-BETA-D-GLUCOSAMINYL-
CC GLYCOPETIDE -> UDP + BETA-D-GALACTOSYL-1,3-N-ACETYL-BETA-D-
CC GLUCOSAMINYLGLYCOPETIDE.
CC -1- PATHWAY: GLYCOSYLATION.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND COLON MUCOSA AND TO A
CC LESSER EXTENT IN COLON ADENOCARCINOMA CELLS. NO EXPRESSION IN
CC HEART, PLACENTA, KIDNEY, LIVER, SKELETAL MUSCLE, LUNG AND
CC PANCREAS.
DR EMBL: AF117222; AAD23451.1; -

```

```
DR EMBL; AB041407; BAA94492.1; -.
DR MIM; 603093; -.
DR InterPro; IPR002659; Galactosyl_T.
DR Pfam; PF01762; Galactosyl_T; 1.
KW Transferase; Glycosyltransferase; Transmembrane; Signal-anchor;
KW Glycoprotein; Golgi stack.
FT DOMAIN 1 4 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 5 20 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DOMAIN 21 326 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 326 AA; 37993 MW; 83271E99B2EE74F5 CRC64;

Query Match 35.4%; Score 587; DB 4; Length 326;
Best Local Similarity 38.6%; Pred. No. 3e-47;
Matches 127; Conservative 52; Mismatches 108; Indels 42; Gaps 7;

QY 12 CLVLGALCLYFSMYSLN-----PFKEQSFYVK--DGN-----FL 45
DB 7 CLYLVTVWCWASALWYLSITRPTSSYTGSKPFSHLTVARKNFTFGNIRPINPHSFEFL 66
QY 46 KLPDTCRQTPPFLVLLVTSSSHQKLAERMAIROTWGMKRMVKGKQLKTFPFLGTTSSAAE 105
DB 67 INEPNCKENIPFLVLIISTHKEFDARQAIRETWGDENNFKGIKTIATFLGKNADPVL 126
QY 106 TKEVDQESQRHGDIIQKDFLVYNYNLTKTMGIEWHRCFQQAFAFMKTDSDMFINVY 165
DB 127 NQWVEQESQIFHDIIVEDFIDSYHNLTGLMGRWVATFCSKAKYVMKTDSDIFVNMND 186
QY 166 LTELKLNKR--TTRFTGFLKNEPIRQPFKSWFVSKSEYPDWDRYPPPCSGTGYVFSG 223
DB 187 LIYKLLKPKTKPRRRYFTGYV--INGGPIRDRVRSKWTMPRDLYPDSNYPFPFCSGTGYFSA 245
QY 224 DVASOVYVSKSVPIKLEDFVGLCLERLNIRLELHLSQTPFFPGG-----LRFVSCLF 278
DB 246 DVAEIYKTSLSHRLHLEDVYVGLCLRLKGIH-----PFQNSGFHNWKMAYSICRY 297
QY 279 RRIVACHFIKPRTLDDYQWALENSRGEDC 307
DB 298 RRVITVHQISPEEMHRIWDMSSKKHLRC 326

RESULT 4
Q9MYM7 PRELIMINARY; PRT; 326 AA.
AC Q9MYM7
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Beta 1,3-galactosyltransferase polypeptide 1.
GN B3GALT1.
OS Pongo pygmaeus (Orangutan),
OS Pan troglodytes (Chimpanzee),
OS Pan paniscus (Pygmy chimpanzee) (Bonobo), and
OS Gorilla gorilla (gorilla)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600, 9598, 9597, 9593;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=P. pygmaeus, P. troglodytes, P. paniscus, and G. gorilla;
RA Liu Y., Saitou N.;
RT "Silver Project.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB041411; BAA94496.1; -.
DR EMBL; AB041408; BAA94493.1; -.
DR EMBL; AB041409; BAA94494.1; -.
DR EMBL; AB041410; BAA94495.1; -.
DR InterPro; IPR002659; Galactosyl_T.
DR Pfam; PF01762; Galactosyl_T; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 326 AA; 37993 MW; 83271E99B2EE74F5 CRC64;
```

```
Query Match 35.4%; Score 587; DB 6; Length 326;
Best Local Similarity 38.6%; Pred. No. 3e-47;
Matches 127; Conservative 52; Mismatches 108; Indels 42; Gaps 7;

QY 12 CLVLGALCLYFSMYSLN-----PFKEQSFYVK--DGN-----FL 45
DB 7 CLYLVTVWCWASALWYLSITRPTSSYTGSKPFSHLTVARKNFTFGNIRPINPHSFEFL 66
QY 46 KLPDTCRQTPPFLVLLVTSSSHQKLAERMAIROTWGMKRMVKGKQLKTFPFLGTTSSAAE 105
DB 67 INEPNCKENIPFLVLIISTHKEFDARQAIRETWGDENNFKGIKTIATFLGKNADPVL 126
QY 106 TKEVDQESQRHGDIIQKDFLVYNYNLTKTMGIEWHRCFQQAFAFMKTDSDMFINVY 165
DB 127 NQWVEQESQIFHDIIVEDFIDSYHNLTGLMGRWVATFCSKAKYVMKTDSDIFVNMND 186
QY 166 LTELKLNKR--TTRFTGFLKNEPIRQPFKSWFVSKSEYPDWDRYPPPCSGTGYVFSG 223
DB 187 LIYKLLKPKTKPRRRYFTGYV--INGGPIRDRVRSKWTMPRDLYPDSNYPFPFCSGTGYFSA 245
QY 224 DVASOVYVSKSVPIKLEDFVGLCLERLNIRLELHLSQTPFFPGG-----LRFVSCLF 278
DB 246 DVAEIYKTSLSHRLHLEDVYVGLCLRLKGIH-----PFQNSGFHNWKMAYSICRY 297
QY 279 RRIVACHFIKPRTLDDYQWALENSRGEDC 307
DB 298 RRVITVHQISPEEMHRIWDMSSKKHLRC 326

RESULT 5
Q920V5 PRELIMINARY; PRT; 305 AA.
AC Q920V5
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase I (Fragment).
GN B3GT1.
OS Mus spicilegus (Steppe mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZBN;
RA Liu Y., Kitano T., Shirosaki T., Moriaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of Five Mus musculus subspecies.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB039143; BAB68667.1; -.
DR InterPro; IPR002659; Galactosyl_T.
DR Pfam; PF01762; Galactosyl_T; 1.
KW Glycosyltransferase; Transferase.
FT NON_TER 1 305
FT NON_TER 305 305
SQ SEQUENCE 305 AA; 35611 MW; A338019F3B06EF5A CRC64;

Query Match 34.8%; Score 576; DB 11; Length 305;
Best Local Similarity 42.3%; Pred. No. 3e-46;
Matches 119; Conservative 43; Mismatches 97; Indels 22; Gaps 5;

QY 23 FMSYSLNPFKEQSFYVKKDGFLKLPDTCRQTPPFLVLLVTSSSHQKLAERMAIROTWGMK 82
DB 39 FGNIRPINPHSF-----EFLNPNCKENIPFLVLIISTHKEFDARQAIRETWGD 92
QY 83 ERNVKQKQLKTFPFLGTTSSAAETKEVDQESQRHGDIIQKDFLVYNYNLTKTMGIEW 142
DB 93 ENNFKGIKIATFLGKNADPVLNQWVEQESQIFHDIIVEDFIDSYHNLTGLMGRWV 152
QY 143 HRCFQQAFAFMKTDSDMFINVYLTLLKLNKR--TTRFTGFLKNEPIRQPFKSWFV 200
DB 153 ATFCSKAKYVMKTDSDIFVNMNDLIYKLLKPKTKPRRRYFTGYV--INGGPIRDRVRSK 211
```

```

QY 201 SKSEYPPDRYPPFCGSGTGYVSGDVASQVYNNVSKSPYKLEDDVFGVGLCLERLNRLEEL 260
Db 212 PRDLYPDSNYPFCGSGTGYVSGDVASQVYNNVSKSPYKLEDDVFGVGLCLERLNRLEEL 267
QY 261 HSQPTFFPGG-----LRFVSVCLFRPRIVACHFIKPRTLDDYW 296
Db 268 ----PFQNSGFNHWKMAISLCRYRVRVTVHQISPEMHRIW 304

RESULT 6
Q91V52
ID O91V52; PRELIMINARY; PRT; 305 AA.
AC O91V52;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE UDP-Gal:betaglcNAC beta 1,3-galactosyltransferase I (Fragment).
GN B3GT1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID-10090;
RN [1]
RP SEQUENCE FROM N.A.;
RC STRAIN-VARIOUS STRAINS;
RA Liu Y., Kitano T., Koide T., Shiroishi T., Moriaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
RT Five Mus musculus subspecies."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB039134; BAB68658.1; -
DR EMBL; AB039135; BAB68659.1; -
DR EMBL; AB039136; BAB68660.1; -
DR EMBL; AB039137; BAB68661.1; -
DR EMBL; AB039138; BAB68662.1; -
DR EMBL; AB039139; BAB68663.1; -
DR EMBL; AB039140; BAB68664.1; -
DR EMBL; AB039141; BAB68665.1; -
DR EMBL; AB039142; BAB68666.1; -
DR InterPro: IPR002659; Galactosyl_T.
DR Pfam: PF01762; Galactosyl_T; 1.
KW Glycosyltransferase; transferase.
FT NON_TER 1
FT NON_TER 305
SQ SEQUENCE 305 AA; 35611 MW; A338019F3B06EF5A CRC64;

Query Match 34.88; Score 576; DB 11; Length 305;
Best Local Similarity 42.38; Pred. No. 3e-46;
Matches 119; Conservative 43; Mismatches 97; Indels 22; Gaps 5;

QY 23 FMSYSLNPKQSPYKKGKGNFLKLPDTCRQTPPFLVLTSSHKOLAERMAIROTGWK 82
Db 39 FGNIRTRPINPSE-----EFLNEPKCKEKNIPFLVILSTHKEFDARQAIRETWGD 92
QY 83 ERMYKQKLTFFLLGTTSSAETKEVDQESORHGDIIQKDFLDVYNNLTAKTMGIW 142
Db 93 ENNPFGKIATFLFLGNADPVLNQWQESQIFHDIIVEDFIDSYHNLTKTLGMRWV 152
QY 143 HRFCPOAAFYKMTDSDMFINDVYLTLLKNNR--TTRFFGFLKLNFFPRQPFKWFV 200
Db 153 ATFCCKAKYVMTDSDIFVNMNDLIYLLKLPSTKPRRYFTGYV-INGGPIRDRSKWYM 211
QY 201 SKSEYPPDRYPPFCGSGTGYVSGDVASQVYNNVSKSPYKLEDDVFGVGLCLERLNRLEEL 260
Db 212 PRDLYPDSNYPFCGSGTGYVSGDVASQVYNNVSKSPYKLEDDVFGVGLCLERLNRLEEL 267
QY 261 HSQPTFFPGG-----LRFVSVCLFRPRIVACHFIKPRTLDDYW 296
Db 268 ----PFQNSGFNHWKMAISLCRYRVRVTVHQISPEMHRIW 304

RESULT 7
O43825
ID O43825 PRELIMINARY; PRT; 422 AA.

```

```

AC 043825;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE UDP-galactose:BETA-N-acetyl-glucosamine-BETA-1,3-galactosyltransferase
DE 2 (EC 2.4.1.-) (BETA3GALT2).
GN GALT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID-9606;
RN [1]
RP SEQUENCE FROM N.A.; FUNCTION, AND TISSUE SPECIFICITY.
RC TISSUE-BRAIN;
RA MEDLINE-98079080; PubMed-9417100;
RX Kolbinger F., Streiff M.B., Katopodis A.G.;
RT "Cloning of a human UDP-galactose:2-acetamido-2-deoxy-D-glucose 3beta-
RT galactosyltransferase catalyzing the formation of type 1 chains."
RL J. Biol. Chem. 273:433-440(1998).
RN [2]
RP SEQUENCE FROM N.A.; FUNCTION, AND TISSUE SPECIFICITY.
RC TISSUE-FETAL BRAIN;
RA Amado M., Almeida R., Carneiro F., Levery S.B., Holmes E.H.,
RA Nomoto M., Hollingsworth M.A., Hassan H., Schwientek T., Nielsen P.A.,
RA Bennett E.P., Clausen H.;
RT "A family of human beta3-galactosyltransferases. Characterization of
RT four members of a UDP-galactose:beta-N-acetyl-glucosamine/beta-
RT nacytyl-galactosamine beta-1,3-galactosyltransferase family."
RL J. Biol. Chem. 273:12770-12778(1998).
CC -1- FUNCTION: THIS PROTEIN IS RESPONSIBLE FOR THE SYNTHESIS OF
CC COMPLEX-TYPE N-LINKED OLIGOSACCHARIDES IN MANY GLYCOPROTEINS A AS
CC WELL AS THE CARBOHYDRATE MOETIES OF GLYCOLIPIDS. GALT1 AND GALT2
CC ENZYMES ACT ON THE SAME ACCEPTOR SUBSTRATES BUT WITH DIFFERENT
CC AFFINITY
CC -1- CATALYTIC ACTIVITY: UDP-GALACTOSE + N-ACETYL-BETA-D-GLUCOSAMINYL-
CC GLYCOSAMINE -> UDP + BETA-D-GALACTOSYL-1,3-N-ACETYL-BETA-D-
CC GLUCOSAMINYLGALACTOSIDE.
CC -1- PATHWAY: GLYCOSYLATION.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART AND TO A LESSER EXTENT IN
CC BRAIN. NO EXPRESSION IN PLACENTA, KIDNEY, LIVER, SKELETAL MUSCLE,
CC LUNG AND PANCREAS.
DR EMBL; Y15014; CAA75245.1; -
DR EMBL; Y15060; CAA75344.1; -
DR MIN; 603018; -
DR InterPro: IPR002659; Galactosyl_T.
DR Pfam: PF01762; Galactosyl_T; 1.
KW Transferase; Glycosyltransferase; Transmembrane; Signal-anchor;
KW Glycoprotein; Golgi stack.
FT DOMAIN 1 24 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 25 43 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DOMAIN 44 422 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 422 AA; 49213 MW; BA62942D6AFB4B0C CRC64;

Query Match 34.38; Score 569; DB 4; Length 422;
Best Local Similarity 41.58; Pred. No. 2.1e-45;
Matches 108; Conservative 53; Mismatches 95; Indels 4; Gaps 2;

QY 52 CRQTPPFLVLTSSHKOLAERMAIROTGWKRMVKGKQLTKFFLLGTT--SSAETKEV 109
Db 146 COEKSPFLILLIAEPGQIARRALROTGWGNSLAPGQITRIFLLGSLIKNGYLQRAI 205
QY 110 DQESORHGDIIQKDFLDVYNNLTAKTMGIWHRFCPOAAFYKMTDSDMFINDVYLT 169
Db 206 LEESRQYHDIQVEYLDYNNLTAKTMGMWVATYCHPIFYVMTKDSDFNTEYLK 265
QY 170 LLKKNRTR--FTFGFLKLNFFPRQPFKWFVSKSEYPPDRYPPFCGSGTGYVSGDVAS 227

```



```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSM/MSE, AND SWN/MSE;
RA Liu Y., Kitano T., Koide T., Shiioishi T., Moriaki K., Saitou N.;
RT *Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
RT Five Mus musculus subspecies*.
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB039148; BAB68672.1; -.
DR EMBL; AB039150; BAB68674.1; -.
DR InterPro; IPR002659; Galactosyl_T.
DR Pfam; PF01762; Galactosyl_T.
KW Glycosyltransferase; Transferase.
FT NON_TER 1 409
FT NON_TER 409 409
SQ SEQUENCE 409 AA; 47206 MW; D7BD514EE40A0157 CRC64;

Query Match 34.2%; Score 567; DB 11; Length 409;
Best Local Similarity 41.2%; Pred. No. 3.1e-45;
Matches 107; Conservative 53; Mismatches 96; Indels 4; Gaps 2;

QY 52 CROTPPFLVLLVTSKHKQLAERMAIRQTWGMKRVKQKLTFFLLGTT--SSAAETKEV 109
Db 140 COEKSPFLILLIAAEFGQIEARRAIRQTWNETLAPGIIIRVFLLGISIKLNGYLQHA 199
QY 110 DOESORHGDIIOKDFLDVYVNLTKTMGIEWHRCPOAAFMKTDSDMFINDVYLTEL 169
Db 200 QESQYHDIIOQEVLDVYVNLTKTMGMNVATYCPHTPYVMKTDSDMFVNTYELIHK 259
QY 170 LLLKNRTTR--FFTGLKNEFFIRQPFKSWFVSKSEYPMWRPFCSTGTVFSGDVAS 227
Db 260 LKLPDLPPRHNYFTGYLMRGYAPNRKDSKWMYMPDLPYSEYRPFVCSGTGYVFGDLAE 319
QY 228 QVNVSKSPYPIKLEDFVGLCLERLNIRLELHSDPTFFPGGLRFSVCLFRIVACHFI 287
Db 320 KIFKYSLGIRRLHLEDVYVGICLAKLRVDPVPPPEFVFNHRVSYSSCKYSLHITSQF 379
QY 288 KPRTLLDYWOALENSRGEDC 307
Db 380 QPSSEIKYWNHLQONKHAC 399

RESULT 14
054905
ID 054905 PRELIMINARY; PRT; 422 AA.
AC 054905;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE UDP-galactose:BETA-N-acetyl-glucosamine-BETA-1,3-galactosyltransferase
DE II (EC 2.4.1.-) (UDP-GAL:BETAGLACNA BETA 1,3-GALACTOSYLTRANSFERASE-II)
DE (BETA3GALT-II).
GN B3GT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC STRAIN=129/SVJ;
RX MEDLINE=98079027; PubMed=9417047;
RA Hennes T., Dinter A., Kuhnert P., Mattu T.S., Rudd P.M., Berger E.G.;
RT "Genomic cloning and expression of three murine UDP-galactose: beta-N-
RT acetylglucosamine beta1,3-galactosyltransferase genes."
RL J. Biol. Chem. 273:58-65(1998).
CC -1- FUNCTION: THIS PROTEIN IS RESPONSIBLE FOR THE SYNTHESIS OF
CC COMPLEX-TYPE N-LINKED OLIGOSACCHARIDES IN MANY GLYCOPROTEINS AS
CC WELL AS THE CARBOHYDRATE MOIETIES OF GLYCOLIPIDS. GALT-I, GALT-II
CC AND GALT-III ENZYMES ACT ON THE SAME ACCEPTOR SUBSTRATES BUT WITH
CC DIFFERENT AFFINITY.
CC -1- CATALYTIC ACTIVITY: UDP-GALACTOSE + N-ACETYL-BETA-D-GLUCOSAMINYL-
CC GLYCOPOLYPEPTIDE -> UDP + BETA-D-GALACTOSYL-1,3-N-ACETYL-BETA-D-
CC GLUCOSAMINYLGALCOPEPTIDE.

```

```

CC -1- COFACTOR: MANGANESE.
CC -1- PATHWAY: GLYCOSYLATION.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND HEART. BARELY DETECTED
CC IN OVARY, COLON AND LYMPH NODES. NO EXPRESSED IN TESTIS, UTERUS,
CC STOMACH, KIDNEY, LIVER, LUNG, BONE MARROW, THYMUS AND SPLEEN.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-14 IS THE INITIATOR.
DR EMBL; AF029791; AAC53524.1; -.
DR MGD; MGI:1349461; B3galt2.
DR InterPro; IPR002659; Galactosyl_T.
DR Pfam; PF01762; Galactosyl_T.
KW Transferase; Glycosyltransferase; Transmembrane; Signal-anchor;
KW Glycoprotein; Golgi stack; Manganese.
FT DOMAIN 1 24
FT TRANSMEM 25 43
FT DOMAIN 44 422
FT CARBOHYD 75 75
FT CARBOHYD 98 98
FT CARBOHYD 119 119
FT CARBOHYD 176 176
FT CARBOHYD 226 226
SQ SEQUENCE 422 AA; 49107 MW; D03D18A7484DCB03 CRC64;

Query Match 34.2%; Score 567; DB 11; Length 422;
Best Local Similarity 41.2%; Pred. No. 3.2e-45;
Matches 107; Conservative 53; Mismatches 96; Indels 4; Gaps 2;

QY 52 CROTPPFLVLLVTSKHKQLAERMAIRQTWGMKRVKQKLTFFLLGTT--SSAAETKEV 109
Db 146 COEKSPFLILLIAAEFGQIEARRAIRQTWNETLAPGIIIRVFLLGISIKLNGYLQHA 205
QY 110 DOESORHGDIIOKDFLDVYVNLTKTMGIEWHRCPOAAFMKTDSDMFINDVYLTEL 169
Db 206 QESQYHDIIOQEVLDVYVNLTKTMGMNVATYCPHTPYVMKTDSDMFVNTYELIHK 265
QY 170 LLLKNRTTR--FFTGLKNEFFIRQPFKSWFVSKSEYPMWRPFCSTGTVFSGDVAS 227
Db 266 LKLPDLPPRHNYFTGYLMRGYAPNRKDSKWMYMPDLPYSEYRPFVCSGTGYVFGDLAE 325
QY 228 QVNVSKSPYPIKLEDFVGLCLERLNIRLELHSDPTFFPGGLRFSVCLFRIVACHFI 287
Db 326 KIFKYSLGIRRLHLEDVYVGICLAKLRVDPVPPPEFVFNHRVSYSSCKYSLHITSQF 385
QY 288 KPRTLLDYWOALENSRGEDC 307
Db 386 QPSSEIKYWNHLQONKHAC 405

RESULT 15
09BZ09
ID 09BZ09 PRELIMINARY; PRT; 422 AA.
AC 09BZ09;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE B3GALT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21218927; PubMed=11318611;
RA Sood R., Bonner T.I., Malakowska I., Stephan D.A., Robbins C.M.,
RA Connors T.D., Morgenbesser S.D., Su K., Faruque M.U., Pinkett H.,
RA Graham C., Baxevanis A.D., Klinger K.W., Landes G.M., Trent J.M.,
RA Carpten J.D.;
RT "Cloning and characterization of 13 novel transcripts and the human
RT RGS8 gene from the 1q25 region encompassing the hereditary prostate
RT cancer (hpc1) locus."
RT Genomics 73:211-222(2001).
DR EMBL; AF288390; AAG0610.1; -.
DR InterPro; IPR002659; Galactosyl_T.

```

DR Pfam; PF01762; Galactosyl_T; 1.
SQ SEQUENCE 422 AA; 49241 MW; BA6294E26AFB4ADC CRC64;
Query Match 34.1%; Score 565; DB 4; Length 422;
Best Local Similarity 41.2%; Pred. No. 4.9e-45;
Matches 107; Conservative 54; Mismatches 95; Indels 4; Gaps 2;
QY 52 CROTPPELVLLVTSQKLAERMAIRQWKGKRMVKGKQLKTFELLGTT--SSAAETKEV 109
DB 146 COEKSPFLILIAAEPGQIEARRAIRRTWGNESLAPGQITRIFFLLGLSIKLNGLYQRAI 205
QY 110 DOESORHGDIIOKDFLDVYVNLTKTMGIEWVHRFCPOAAAFVMTKDSDFINVDYLTTEL 169
DB 206 LEESRQYHDIIOQYLDYTYVNLTKTMGMWVATYCPHIPYVMTKDSDFVNTTEYLIN 265
QY 170 LLKKNRTTR--FFTGFELKNEFFIRQPFESKWFVSKSEYPMDRPPFCSGTGYVFGDVAS 227
DB 266 LLKPDLPFRHNYFTCYLMRGVAPNRKDSKWMPDLYPSERYPVFCSTGYVFGDLAE 325
QY 228 QVYNVSKSVPIKLEDVFGVGLERLNRLEELHSQPTFFPGGLRFSVCLFRRIYACHFI 287
DB 326 KIFKVSGLGIRLHLEDVVGICLAKLRIDPVPPEFVFNHWRVSYSSCKYSHLITSHQF 385
QY 288 KPTLLDYWOALENSRGEDC 307
DB 386 QPSELIKYWNHLOQKNHAC 405

Search completed: April 14, 2003, 12:09:05
Job time : 35 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 14:50:38 ; Search time 154.708 Seconds
(without alignments)
13552.085 Million cell updates/sec

Title: US-09-914-152-2_COPY_401_1331
Perfect score: 931
Sequence: 1 aatggctttccgaagatga.....gggaagattgtccgcctgtc 931

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*

1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*

2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*

3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*

4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*

5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*

6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*

7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*

8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*

9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*

10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*

11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*

12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*

13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*

14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*

15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*

16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*

17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*

18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*

19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*

20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*

21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*

22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*

23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*

24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 931 | 100.0 | 2775 | 21 | Human beta-1,3 gal |
| 2 | 930 | 99.9 | 10562 | 21 | Human beta3Gal-T5 |
| 3 | 925.2 | 99.4 | 1011 | 21 | Human beta3Gal-T5 |
| 4 | 163.4 | 17.6 | 1739 | 15 | AAQ67067 |
| 5 | 130.4 | 14.0 | 1773 | 20 | Beta-1,3-galactosy |
| 6 | 130.4 | 14.0 | 1697 | 24 | Human Dendriac cDN |
| 7 | 130.4 | 14.0 | 2095 | 21 | Human cDNA encodin |
| 8 | 130.4 | 14.0 | 2095 | 22 | Membrane-bound pro |
| 9 | 130.4 | 14.0 | 2095 | 22 | Human DNA encoding |
| | | | | | Human PRO1074 cDNA |

| | | | | | | |
|----|-------|------|------|----|----------|--------------------|
| 10 | 130.4 | 14.0 | 2095 | 22 | AAF44168 | Human PRO1074 (UNO |
| 11 | 130.4 | 14.0 | 2168 | 22 | AAH15711 | Human cDNA sequenc |
| 12 | 130.4 | 14.0 | 2189 | 20 | AAH87192 | Human Dendriac cDN |
| 13 | 122.4 | 13.1 | 1266 | 20 | AAH35710 | cDNA encoding a pr |
| 14 | 122.4 | 13.1 | 2420 | 20 | AAH35711 | cDNA encoding a pr |
| 15 | 114.8 | 12.3 | 1134 | 21 | AAH8791 | DNA encoding the b |
| 16 | 108.2 | 11.6 | 1134 | 22 | AAH09356 | Human beta-1,3-gal |
| 17 | 93.2 | 10.0 | 1725 | 23 | ABL14319 | Drosophila melanog |
| 18 | 93.2 | 10.0 | 1725 | 23 | ABL14319 | Drosophila melanog |
| 19 | 89.8 | 9.6 | 1116 | 19 | AAV49599 | Human epidermoid c |
| 20 | 89.8 | 9.6 | 1208 | 24 | AAH16945 | Human beta 1,3-N-a |
| 21 | 89.8 | 9.6 | 2180 | 22 | AAH29257 | Human beta 1,3-N-a |
| 22 | 89.8 | 9.6 | 2186 | 19 | AAV49598 | Human epidermoid c |
| 23 | 89.8 | 9.6 | 2186 | 24 | ABL14319 | Human epidermoid c |
| 24 | 89.8 | 9.6 | 2186 | 22 | AAH29257 | Nucleotide sequenc |
| 25 | 89.8 | 9.6 | 2205 | 22 | AAH29256 | Human cDNA encodin |
| 26 | 89.8 | 9.6 | 2210 | 22 | AAH46068 | Human beta 1,3-N-a |
| 27 | 89.8 | 9.6 | 2427 | 23 | ABV22526 | Human DNA encoding |
| 28 | 89.8 | 9.6 | 2427 | 23 | ABV28341 | Human prostate exp |
| 29 | 87.6 | 9.4 | 1337 | 24 | AAH16946 | Murine beta1,3-N-a |
| 30 | 83.4 | 9.0 | 3033 | 23 | ABL14314 | Drosophila melanog |
| 31 | 83.4 | 9.0 | 3330 | 23 | ABL14322 | Drosophila melanog |
| 32 | 81.8 | 8.8 | 1271 | 20 | AAH97916 | Human secreted pro |
| 33 | 81.8 | 8.8 | 1271 | 20 | AAH87194 | Human Brainiac-3 c |
| 34 | 78.4 | 8.4 | 1296 | 22 | AAH29258 | Human beta 1,3-N-a |
| 35 | 78.4 | 8.4 | 1420 | 21 | AAH58790 | DNA encoding the b |
| 36 | 78.4 | 8.4 | 1642 | 21 | AAH58115 | Human PRO4344 nucl |
| 37 | 78.4 | 8.4 | 1643 | 24 | ABK69978 | cDNA encoding huma |
| 38 | 77.8 | 8.4 | 1359 | 23 | AAH84438 | DNA encoding novel |
| 39 | 77.8 | 8.4 | 1360 | 24 | AAH35225 | Human TRNFR-10 cDN |
| 40 | 74.6 | 8.0 | 977 | 21 | AAH57383 | DNA encoding a Bra |
| 41 | 74.6 | 8.0 | 2117 | 24 | AAH35219 | Human TRNFR-4 cDNA |
| 42 | 74.6 | 8.0 | 2684 | 21 | AAH58117 | Human PRO4397 nucl |
| 43 | 74.6 | 8.0 | 2684 | 24 | ABK69977 | cDNA encoding huma |
| 44 | 70.2 | 7.5 | 400 | 21 | AAH94080 | Haematopoietic ste |
| 45 | 68.6 | 7.4 | 1191 | 21 | AAH53023 | Human beta-1,3-gal |

ALIGNMENTS

RESULT 1
AAA93875
ID AAA93875 standard; DNA; 2775 BP.
XX
XX AAA93875;
AC
XX
XX 15-JAN-2001 (first entry)
XX
XX Human beta-1,3 galactose transferase encoding DNA.
XX
XX Beta-1,3 galactose transferase; treatment; diagnosis; cancer; human;
XX digestive system; ds.
XX Homo sapiens.
XX
XX WO2000050608-A1.
XX
XX 31-AUG-2000.
XX
XX 24-FEB-2000; 2000WO-JP01070.
XX
XX 25-FEB-1999; 99JP-0047571.
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX
XX Narimatsu H, Isshiki S, Togayachi A, Sasaki K;
XX
XX WPI; 2000-549409/50.
XX
XX P-PSDB; AAB93875.
XX
XX Beta-1,3 galactose transferase and DNA encoding it, useful for
XX synthesis of type 1 sialyl Lewis, a carbohydrate for treatment of

PT digestive system cancer -
PS Claim 5; Page 99-102; 123pp; Japanese.
XX This invention relates to a polypeptide (I) with beta-1,3 galactose
CC transferase activity, or variants of (I) comprising amino acid additions,
CC deletions and/or substitutions. Included in the invention is DNA encoding
CC all or part of (I); expression vectors containing the DNA, host cells
CC transformed by the vectors; a method for the preparation of the
CC polypeptide by culture of the transformants or by expression in the milk
CC of a transgenic mammal, and antibodies recognising (I). The Beta-1,3
CC galactose transferase protein transfers galactose by beta-1,3 bonding to
CC N-acetylglucosamine present in a non-cyclic carbohydrate chain (such as
CC GlcNAc-beta1-3gal-beta1-4Glc) to give Gal-beta1-3GlcNAc. The protein and
CC DNA encoding it are useful for the treatment and diagnosis of cancer of
CC the digestive system. The present sequence represents Beta-1,3 galactose
CC transferase encoding DNA.
XX
SQ Sequence 2775 BP; 681 A; 698 C; 669 G; 727 T; 0 other;
Query Match 100.0%; Score 931; DB 21; Length 2775;
Best Local Similarity 100.0%; Pred. No. 6.7e-278;
Matches 931; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AATGGCTTTCCGAGATGAGATTGATGTATATTTGCCCTTCGTGTTCTGGGGCTCTTTG 60
DB 401 AATGGCTTTCCGAGATGAGATTGATGTATATTTGCCCTTCGTGTTCTGGGGCTCTTTG 460
QY 61 TTTGTATTTTACATGTACAGTCTAAATCCTTTCAAGAACAGTCTCTTTGTACAGAA 120
DB 461 TTTGTATTTTACATGTACAGTCTAAATCCTTTCAAGAACAGTCTCTTTGTACAGAA 520
QY 121 AGACGGGAATTCCTTAAGCTCCAGATACAGACTGCAGCGACAGACCTCCCTTCCTCGT 180
DB 521 AGACGGGAATTCCTTAAGCTCCAGATACAGACTGCAGCGACAGACCTCCCTTCCTCGT 580
QY 181 CCTGCTGGTACCTCATCCCAACACAGTTGGCTGAGCGCATGGCCATCCGGCAGACGTG 240
DB 581 CCTGCTGGTACCTCATCCCAACACAGTTGGCTGAGCGCATGGCCATCCGGCAGACGTG 640
QY 241 GGGGAAAGAGAGGATGGTGAAGGGAAGCAGCTGAAGACATTCTTCTCCTCGGGACAC 300
DB 641 GGGGAAAGAGAGGATGGTGAAGGGAAGCAGCTGAAGACATTCTTCTCCTCGGGACAC 700
QY 301 CAGCAGTGCAGCGAAACAAAGAGGTGGACCGAGAGACCCAGCGACACGATTTAT 360
DB 701 CAGCAGTGCAGCGAAACAAAGAGGTGGACCGAGAGACCCAGCGACACGATTTAT 760
QY 361 CCAGAGGATTTCTTAGAGCTTATTACAACTGACCTGACCCGAAAGACCATGCGGATAGA 420
DB 761 CCAGAGGATTTCTTAGAGCTTATTACAACTGACCTGACCCGAAAGACCATGCGGATAGA 820
QY 421 ATGGGTCCATCGCTTTGTCTCAGCGCGGCTTTGTGATGAAACAGACTCAGACATGTT 480
DB 821 ATGGGTCCATCGCTTTGTCTCAGCGCGGCTTTGTGATGAAACAGACTCAGACATGTT 880
QY 481 CAICAAATGTTGACTATCGACTGACTGCTTCTGAAGAAAAACAGAACACCGAGTTT 540
DB 881 CATCAATGTTGACTATCGACTGACTGCTTCTGAAGAAAAACAGAACACCGAGTTT 940
QY 541 CACTGGCTTCTGAACCTCAATGAGTTTCCCATCAGGAGCGCATTCAGCACTGTTTCT 600
DB 941 CACTGGCTTCTGAACCTCAATGAGTTTCCCATCAGGAGCGCATTCAGCACTGTTTCT 1000
QY 601 CAGTAAATCTGAATTCGTTGGACAGGTACCACCATTTCTCTCGGCACCGGCTACGT 660
DB 1001 CAGTAAATCTGAATTCGTTGGACAGGTACCACCATTTCTCTCGGCACCGGCTACGT 1060
QY 661 GTTTTCTGGCAGCTGGGAGTCAAGTGTACAAATGCTCTCAAGAGCGTCCCATACATTAA 720
DB 1061 GTTTTCTGGCAGCTGGGAGTCAAGTGTACAAATGCTCTCAAGAGCGTCCCATACATTAA 1120
QY 721 ACTGGAAGACGTTTGTGGGCTCTGCTCGAAGAGGCTGAACATCAGATTGGAGGAGCT 780

DB 1121 ACTGGAAGACGTTGTTGGGGCTCTGCCTCGAAAGGCTGAACATCAGATTGGAGGAGCT 1180
QY 781 CCACTCCAGCCGACCTTTTTCAGGGGCTTACGCTTCTCGTATGCTCTTCAGGAG 840
DB 1181 CCACTCCAGCCGACCTTTTTCAGGGGCTTACGCTTCTCGTATGCTCTTCAGGAG 1240
QY 841 GATCGTGGCTGCTCATCATCAAGCTCGGACTCTCTTGGACTACTGGCAGGCTCTAGA 900
DB 1241 GATCGTGGCTGCTCATCATCAAGCTCGGACTCTCTTGGACTACTGGCAGGCTCTAGA 1300
QY 901 GAATTCGGGGGGAAGATTGTCGCCCTGTC 931
DB 1301 GAATTCGGGGGGAAGATTGTCGCCCTGTC 1331
RESULT 2
AAA93876
ID AAA93876 standard; DNA; 10562 BP.
XX
AC AAA93876;
XX
DT 15-JAN-2001 (first entry)
XX
DE Human beta3Gal-T5 encoding DNA.
XX
KW Beta-1,3 galactose transferase; treatment; diagnosis; cancer; human;
KW digestive system; beta3Gal-T5; ds.
XX
OS Homo sapiens.
XX
PN WO2000050608-A1.
XX
PD 31-AUG-2000.
XX
PF 24-FEB-2000; 2000WO-JP01070.
XX
PR 25-FEB-1999; 99JP-0047571.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Narimatsu H, Isshiki S, Togayachi A, Sasaki K;
XX
DR WPI; 2000-549409/50.
XX
PT Beta-1,3 galactose transferase and DNA encoding it, useful for
PT synthesis of type 1 sialyl Lewis, a carbohydrate for treatment of
PT digestive system cancer -
XX
PS Claim 31; Page 103-111; 123pp; Japanese.
XX
CC This invention relates to a polypeptide (I) with beta-1,3 galactose
CC transferase activity, or variants of (I) comprising amino acid additions,
CC deletions and/or substitutions. Included in the invention is DNA encoding
CC all or part of (I); expression vectors containing the DNA, host cells
CC transformed by the vectors; a method for the preparation of the
CC polypeptide by culture of the transformants or by expression in the milk
CC of a transgenic mammal, and antibodies recognising (I). The Beta-1,3
CC galactose transferase protein transfers galactose by beta-1,3 bonding to
CC N-acetylglucosamine present in a non-cyclic carbohydrate chain (such as
CC GlcNAc-beta1-3gal-beta1-4Glc) to give Gal-beta1-3GlcNAc. The protein and
CC DNA encoding it are useful for the treatment and diagnosis of cancer of
CC the digestive system. The present sequence represents a Beta3Gal-T5
CC encoding DNA sequence.
XX
SQ Sequence 10562 BP; 2610 A; 2415 C; 2574 G; 2963 T; 0 other;
Query Match 99.9%; Score 930; DB 21; Length 10562;
Best Local Similarity 100.0%; Pred. No. 2.9e-277;
Matches 930; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ATGGCTTCCCCAAGATGAGATTGATGTATTTGCTTCTGGTTCGGGGCTCTTTGT 61

Db 8234 ATGGCTTTCCCGAAGATGAGATTGATGTATATTTGCTTCTGTTCTGGGGCTCTTTGT 8293
Qy 62 TTGTATTTTACATCTACAGTCTAAATCTTTCAAGAACAGTCTCTTTGTTTACAGAAA 121
Db 8294 TTGTATTTTACATCTACAGTCTAAATCTTTCAAGAACAGTCTCTTTGTTTACAGAAA 8353
Qy 122 GACGGGAATCTTCTTAAGCTCCAGATACAGACTGCAGGCAGACACCTCCCTTCCTCGTC 181
Db 8354 GACGGGAATCTTCTTAAGCTCCAGATACAGACTGCAGGCAGACACCTCCCTTCCTCGTC 8413
Qy 182 CTGCTGGTGACCTCATCCCAACAGTTGGCTGAGCGATGGCCATCCGGCAGACGTGG 241
Db 8414 CTGCTGGTGACCTCATCCCAACAGTTGGCTGAGCGATGGCCATCCGGCAGACGTGG 8473
Qy 242 GGAAGAAGAGAGATCGTGAAGGGAAGCAGCTGAAGACATTTCTCTCTGGGGACCA 301
Db 8474 GGAAGAAGAGAGATCGTGAAGGGAAGCAGCTGAAGACATTTCTCTCTGGGGACCA 8533
Qy 302 AGCAGTGCAGCGGAACAAAGAGGTGGACCAAGGAGCCAGCGACAGCGGGGACATATTC 361
Db 8534 AGCAGTGCAGCGGAACAAAGAGGTGGACCAAGGAGCCAGCGACAGCGGGGACATATTC 8593
Qy 362 CAGAAGGATTTCTAGACGTCTATTACAACTGACCCCTGAAGACCATGATGGGCATAGAA 421
Db 8594 CAGAAGGATTTCTAGACGTCTATTACAACTGACCCCTGAAGACCATGATGGGCATAGAA 8653
Qy 422 TGGGTCCATCGCTTTGTCTCAGCGGGCTTTGTGATGATGAACAAACAGACTCAGACATGTC 481
Db 8654 TGGGTCCATCGCTTTGTCTCAGCGGGCTTTGTGATGATGAACAAACAGACTCAGACATGTC 8713
Qy 482 ATCAATGTTGACTATCTGACTGAATCTCTCTGAAAGAAAACAGAACCAACAGGTTTTC 541
Db 8714 ATCAATGTTGACTATCTGACTGAATCTCTCTGAAAGAAAACAGAACCAACAGGTTTTC 8773
Qy 542 ACTGGCTTCTTGAACCTCAATGAGTTTCCATCAGGAGCCATTCAGCAAGTGGTTGTC 601
Db 8774 ACTGGCTTCTTGAACCTCAATGAGTTTCCATCAGGAGCCATTCAGCAAGTGGTTGTC 8833
Qy 602 AGTAATCTGAATATCCGTGGGACAGGTACCACCATCTCTCGCGCACCGGCTACGTG 661
Db 8834 AGTAATCTGAATATCCGTGGGACAGGTACCACCATCTCTCGCGCACCGGCTACGTG 8893
Qy 662 TTTTCTGGCGAGTGGCGAGTCAGTGTACAAATGTCTCCAAAGAGGCTCCCATACATTA 721
Db 8894 TTTTCTGGCGAGTGGCGAGTCAGTGTACAAATGTCTCCAAAGAGGCTCCCATACATTA 8953
Qy 722 CTGGAGACGTGTTGTGGGCTCTGCTCCAAAGGCTGAACATCAGATTGGAGGAGCTC 781
Db 8954 CTGGAGACGTGTTGTGGGCTCTGCTCCAAAGGCTGAACATCAGATTGGAGGAGCTC 9013
Qy 782 CACTCCCGAGCGGACCTTTTCCAGGGGCTTACGCTTCTCCGATGCTTTCAGGAGG 841
Db 9014 CACTCCCGAGCGGACCTTTTCCAGGGGCTTACGCTTCTCCGATGCTTTCAGGAGG 9073
Qy 842 ATCGTGCCCTGCCATCTCATCAAGCCTCGGACTCTCTTTGGACTACTGGCAGGCTCTAG 901
Db 9074 ATCGTGCCCTGCCATCTCATCAAGCCTCGGACTCTCTTTGGACTACTGGCAGGCTCTAG 9133
Qy 902 AATTCGGGGGGAAGATTGTCGCTGTC 931
Db 9134 AATTCGGGGGGAAGATTGTCGCTGTC 9163

RESULT 3

AAA27959

ID AAA27959 standard; DNA: 1011 BP.

XX

AC AAA27959;

XX

DT 15-AUG-2000 (first entry)

XX

DE Human beta3Gal-T5 gene sequence.

XX

XX

XX

XX

XX

XX

KW UDP-D-galactose; beta-N-acetylglucosamine beta1,3-galactosyltransferase;
KW Beta3Gal-T5; ss; human; chromosome 21q22.3; galactosylation;
KW beta1,3-galactosyl glycosylated saccharide production; glycopeptide;
glycoprotein.
XX

OS Homo sapiens.

XX Key Location/Qualifiers
CDS 79..1011
FT /*tag= a
FT /*product= "Beta3Gal-T5"
FT complement (79..98)
FT /*tag= b
FT primer_bind complement (150..170)
FT /*tag= c
FT primer_bind 991..1011
FT /*tag= d

XX WO200029558-A1.

XX 25-MAY-2000.

XX 11-NOV-1999; 99WO-US26807.

XX 13-NOV-1998; 98DK-0001483.

XX (CLAU/) CLAUSEN H.

XX Clausen H, Amado M;

XX WPI: 2000-399728/34.

XX P-PSDB: AAY94641.

XX Novel nucleic acid sequence encoding human

UDP-galactose:beta-N-acetylglucosamine beta1,3-galactosyltransferase
useful for obtaining beta1,3-galactosyl glycosylated saccharides and
glycopeptides or glycoproteins

XX Claim 7; Fig 1; 74pp; English.

XX The present invention relates to a nucleic acid sequence encoding

UDP-D-galactose:beta-N-acetylglucosamine beta1,3-galactosyltransferase
(beta3Gal-T5). Beta3 transferases add galactose to the hydroxy group at
carbon 3 of 2-acetamido-2-deoxy-D-glucose (GlcNAc). The present sequence
represents the human beta3Gal-T5 gene sequence. The beta3Gal-T5 gene is
located on human chromosome 21q22.3. Beta3Gal-T5 is a type II

transmembrane glycoprotein. The invention also relates to the beta3Gal-T5

protein sequence, a nucleic acid vector comprising the beta3Gal-T5

nucleotide sequence, a host cell comprising the vector, and a method for

the production of the beta3Gal-T5 protein from the host cells. The

methods of the invention can be used for recombinant production of

beta3Gal-T5 for use as a catalyst and for recombinant production of

peptides or proteins with appropriate galactosylation. The beta3Gal-T5

protein can be used to obtain beta1,3-galactosyl glycosylated

saccharides, glycopeptides or glycoproteins.

XX Sequence 1011 BP; 247 A; 256 C; 251 G; 257 T; 0 other;

Query Match

Best Local Similarity 99.4%; Score 925.2; DB 21; Length 1011;

Matches 927; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ATGGCTTTCCCGAAGATGAGATTGATGTATATTTGCTTCTGTTCTGGGGCTCTTTGT 61

Db 79 ATGGCTTTCCCGAAGATGAGATTGATGTATATCTGCTTCTGGTCTGGGGCTCTTTGT 138

Qy 62 TTGTATTTTACATCTACAGTCTAAATCTTTCAAGAACAGTCTCTTTGTTTACAGAAA 121

Db 139 TTGTATTTTACATCTACAGTCTAAATCTTTCAAGAACAGTCTCTTTGTTTACAGAAA 198

Qy 122 GACGGGAATCTTCTTAAGCTCCAGATACAGACTGCAGGCAGACACCTCCCTTCCTCGTC 181

Db 199 GACGGGAATCTTCTTAAGCTCCAGATACAGACTGCAGGCAGACACCTCCCTTCCTCGTC 258

| | | | |
|----|-----|--|------|
| QY | 182 | CTGCTGGTGACCTTCATCCCAAAACAGTTGGCTGAGCGCATGGCCATCCGGCAGACGTGG | 241 |
| | | | |
| Db | 259 | CTGCTGGTGACCTTCATCCCAAAACAGTTGGCTGAGCGCATGGCCATCCGGCAGACGTGG | 318 |
| | | | |
| QY | 242 | GGGAAGAGAGAGTGGTGAAGGGAAACAGCTGAAGACATCTTCTCTCTGGGGACCAACC | 301 |
| | | | |
| Db | 319 | GGGAAGAGAGAGCGTGAAGGAAAGAGCTGAAGACATCTTCTCTCTGGGGACCAACC | 378 |
| | | | |
| QY | 302 | AGCAGTGCACGGAAACGAAGAGGTGGACAGGAGACGCGACACGGGACATTTATC | 361 |
| | | | |
| Db | 379 | AGCAGTGCACGGAAACGAAGAGGTGGACAGGAGACGCGACACGGGACATTTATC | 438 |
| | | | |
| QY | 362 | CAGAAGGATTTCTAGACGCTCTATTACAATCTGACCTCAAGACCATGATGGGCATAGAA | 421 |
| | | | |
| Db | 439 | CAGAAGGATTTCTAGACGCTCTATTACAATCTGACCTCAAGACCATGATGGGCATAGAA | 498 |
| | | | |
| QY | 422 | TGGTCCATCGCTTTTGTCTCAGGCGGGCTTGTGATGAAACACAGCTCAGACATGTTTC | 481 |
| | | | |
| Db | 499 | TGGTCCATCGCTTTTGTCTCAGGCGGGCTTGTGATGAAACACAGCTCAGACATGTTTC | 558 |
| | | | |
| QY | 482 | ATCAATGTTGACTATCTGACTGAATGCTTCTGAGAGAAAACAGAACCAACCAAGTTTTC | 541 |
| | | | |
| Db | 559 | ATCAATGTTGACTATCTGACTGAATGCTTCTGAGAGAAAACAGAACCAACCAAGTTTTC | 618 |
| | | | |
| QY | 542 | ACTGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGCAGCCATTCAGCAAGTGGTTTGC | 601 |
| | | | |
| Db | 619 | ACTGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGCAGCCATTCAGCAAGTGGTTTGC | 678 |
| | | | |
| QY | 602 | AGTAAATCTGAATATCCGTGGGACAGGTACCCACCATCTGCTCCGGCACCGGCTACGCTG | 661 |
| | | | |
| Db | 679 | AGTAAATCTGAATATCCGTGGGACAGGTACCCACCATCTGCTCCGGCACCGGCTACGCTG | 738 |
| | | | |
| QY | 562 | TTTTCCTGGCACGTGGCGAGTCAGGTGTACAATGTCTCCAAGAGGCTGCCATACATTTAA | 721 |
| | | | |
| Db | 739 | TTTTCCTGGCACGTGGCGAGTCAGGTGTACAATGTCTCCAAGAGGCTGCCATACATTTAA | 798 |
| | | | |
| QY | 722 | CTGGAAGACGTGTTTGTGGGGCTCTGCCCTCGAAAGGCTGAACATCAGATTGGAGAGCTC | 781 |
| | | | |
| Db | 799 | CTGGAAGACGTGTTTGTGGGGCTCTGCCCTCGAAAGGCTGAACATCAGATTGGAGAGCTC | 858 |
| | | | |
| QY | 782 | CACTCCGACCGACCTTTTTCAGGGGGCTTACGGTTCTCCGTATGCGCTTTCAGGAGG | 841 |
| | | | |
| Db | 859 | CACTCCGACCGACCTTTTTCAGGGGGCTTACGGTTCTCCGTATGCGCTTTCAGGAGG | 918 |
| | | | |
| QY | 842 | ATCGTGGCTTGCACATTCATCAAGCCTCGGACTCTCTTGGACTACTGGCAGGCTCTAGAG | 901 |
| | | | |
| Db | 919 | ATCGTGGCTTGCACATTCATCAAGCCTCGGACTCTCTTGGACTACTGGCAGGCTCTAGAG | 978 |
| | | | |
| QY | 902 | AATTCCGGGGGAAGATTGTCCGCGTGC | 931 |
| | | | |
| Db | 979 | AATTCCGGGGGAAGATTGTCCGCGTGC | 1008 |
| | | | |

RESULT 4
AAQ67067
ID AAQ67067 standard; cDNA to mRNA; 1739 BP.

AAC 67067;

DT 14--MAR-1995 (first entry)

Beta-1,3-galactosyltransferase cDNA.

Beta-1,3-galactosyltransferase; enzyme; saccharide chain; pAMOPRW1;
KJM-1 cells; ds.

Homo sapiens.

| | | | |
|----|-----|---------------------|--|
| AA | | | |
| FH | Key | Location/Qualifiers | |
| FT | CDS | 676..1656 | |
| FT | | /tag= a | |
| FT | | /product= Beta-1,3- | |

XX PN JP06181759-A.
XX PD 05-JUL-1994.
XX PF 16-DEC-1992; 92JP-0336436.
XX PR 16-DEC-1992; 92JP-0336436.
XX PA (KYOW) KYOWA HAKKO KOGYO KK.
XX DR WPI; 1994-251683/31.
XX DR P-PSDB; AAR57433.
XX PT Beta-galactosyl-transferase DNA and protein - useful for prodn. of
XX PT saccharide chains
XX Claim 3; Page 22-24; 47pp; Japanese.
XX This sequence encodes a beta-1,3-galactosyltransferase. This enzyme
CC can be used to produce physiologically active saccharide chains and
CC variants, and for improvement of saccharide chains bound to
CC physiologically active proteins. This cDNA represents a fragment
CC of the plasmid pAMOPRWMI which was cloned in KJM-1 cells.
XX SQ Sequence 1739 BP; 473 A; 407 C; 422 G; 437 T; 0 other;

Query Match 17.68; Score 163.4; DB 15; Length 1739;
Best Local Similarity 55.78; Pred. No. 8.4e-40;
Matches 358; Conservative 0; Mismatches 276; Indels 9; Gaps 2;

Qy 129 ACTTCTTAAGCTCCCAGATACAGACTGCAGGCGAGACACCTCCTTCCTCGTGCTGGT 188
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 866 AATTTCCTTAAACGAGGCCCAATAAATGTGAGAAAAACATTCTTTCTGTTATCTCA 925
Qy 189 TGACCTCATCCACAACAGTTGGCTGAGCGCATGGCCATCCGCGACAGCTGGGGGAAG 248
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 926 TCAGCACCACTCAGCAAGAATTTGATCCCGTCAGGCAATCAGAGACAGCTGGGGGATG 985
Qy 249 AGAGGATGGTGAAGGGAAAGCAGCTGAAGACATTCTTCTCTGGGACCACAGCAGTG 308
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 986 AGAACAACTTTAAGGGATCAAGATAGCCACCTGTTCTCTGGGCAAGATCTGATC 1045
Qy 309 CAGCGGAACAAAGAGGTGGACGAGAGCCACGACAGCGGGGACATTATCCAGAAG 368
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1046 CTGTTCTCAATCAGATGGTGGAGCAAGAGGCCAAATCTTCCATGATATCATCGTG 1105
Qy 369 ATTTCCTAGACGCTATTACAACTCGACCTGGAAGACCATGATGGGCATAGATGGTCC 428
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1106 ACTTTATGACTCCTACCATACCTTACCTCAAAACATTANTGGGGATGAGATGGGTG 1165
Qy 429 ATCCGTTTTGTCFCAGGGCGGCTTTGTATGAAAACAGACTCAGACATGTTCAATCA 488
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1166 CCACCTTTGTTCAAAAGCCAAAGTATGTCATGAAAACAGACAGCGCATTTTGTAAACA 1225
Qy 489 TTGACTATCTGACTGAACCTGCTCTGAA-----GAAAAACAGAACACCAAGGTTTCA 542
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1226 TGGACATCTATTATTAATTAATCTGAAACCCCTCCACCAAGCCACGAAGAGGATTTTA 1285
Qy 543 CTGCTCTTTGAAACTCAATGAGTTTCCCCATCAGGAGCCATTACGCAAGTGGTTGTCA 602
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1286 CTGCTATGTCA---TTAATGGGAGCCGATTCGGGATGTCGCCAGTAAAGTATATGC 1342
Qy 603 GTAATCTGAATATCCGTGGGACAGGTACCCACCATCTGCTCCGGACCGGTACGTGT 662
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1343 CCAGGGATTTGTACCCACAGACAGTAACCTACCCACTTTCTGTCGGGACTGTACATCT 1402
Qy 663 TTTCTGCGCAGCTGGCGAGTCAGTGTACAAATGCTCCAAGACGCTCCCAACATTAAC 722
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1403 TTTCAGCCGATGTAGTGAACCTCATTTACAAGACCTTCACTCCACACAGGCTGCTTACC 1462
Qy 723 TGGAAAGACGTTTGTGGGCTCTGCTCGAAAGCGCTGAACAT 765
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

| | | | | |
|-----------------------|---|--------------------|-----------|--------------|
| Query Match | 14.08; | Score 130.4; | DB 20; | Length 1773; |
| Best Local Similarity | 48.9%; | Pred. No. 1.5e-29; | | |
| Matches 387; | Conservative 0; | Mismatches 396; | Indels 9; | Gaps 1; |
| QY 111 | TTTACAGAAAGACGGGAACCTTCCTTAAGCTCCAGATACAGCTGCAGGACAGACACCTC | 170 | | |
| | | | | |

| | | |
|----|--|--|
| FT | CDS | /note= "This region is specifically claimed in claim 1" |
| FT | FT | 109..1104 |
| FT | FT | /*tag= b |
| FT | FT | /product= "Beta1,3-acetylgalactosamine transferase" |
| FT | polyA_signal | 1157..1162 |
| FT | FT | /*tag= c |
| XX | | |
| PN | JP2002085069-A. | |
| XX | | |
| PD | 26-MAR-2002. | |
| XX | | |
| PX | 08-SEP-2000; 2000JP-0273835. | |
| PF | | |
| XX | 08-SEP-2000; 2000JP-0273835. | |
| PR | (SEKG) SEIKAGAKU KOGYO CO LTD. | |
| XX | (FURU/) FURUKAWA K. | |
| XX | | |
| DR | WPI: 2002-378274/41. | |
| XX | | |
| DR | P-PSDB; AAU80224. | |
| PT | | |
| XX | Preparation of beta-1,3-acetylgalactosamine transferase - | |
| PS | Claim 1; Page 11-13; 15pp; Japanese. | |
| XX | | |
| CC | The invention relates to the preparation of beta-1,3-acetylgalactosamine | |
| CC | (a Gb4 sugar chain molecule) comprising introducing a DNA appearing as | |
| CC | ABK51201 (S1) or a DNA hybridising with S1 or a base sequence | |
| CC | complementary to the base sequence or part of these base sequences under | |
| CC | a stringent condition into a cell and growing the cell to express beta-1, | |
| CC | 3-acetylgalactosamine transferase and collecting the cell to express beta-1, | |
| CC | 3-acetylgalactosamine transferase and collecting it. Also included is the | |
| CC | preparation of Gb4 sugar chain comprising contacting a polypeptide | |
| CC | comprising the beta-1,3-acetylgalactosamine transferase protein appearing | |
| CC | as AAU80224 (A) or a polypeptide consisting of an amino acid sequence in | |
| CC | which at least one amino acid is replaced, deleted, inserted or | |
| CC | transferred in the amino acid sequence (A) and having enzymatic activity | |
| CC | transferring N-acetylgalactosamine residue from an N-acetylgalactosamine | |
| CC | donor to the C3 site of the galactose residue in Gb3 sugar chain which is | |
| CC | the receptor with an N-acetylgalactosamine donor and Gb3 sugar chain. | |
| CC | The present sequence is the cDNA encoding beta-1,3-acetylgalactosamine | |
| CC | transferase. | |
| XX | | |
| SQ | Sequence 1897 BP; 596 A; 327 C; 366 G; 508 T; 0 other; | |
| | | |
| | Query Match | 14.0%; Score 130.4; DB 24; Length 1897; |
| | Best Local Similarity | 48.9%; Pred. No. 1.5e-29; |
| | Matches 387; Conservative | 0; Mismatches 396; Indels 9; Gaps 1 |
| QY | 111 | TTTACAGAAGACGGGNACTTCCTTAGCTCCCAGATACACACTGCAGGCGAGACACCTC 170 |
| Db | 281 | TTTTACACAAGACATTTCACATTCGAGAGCATTCAAACCTGCTCATCAAAATC 340 |
| QY | 171 | CCTTCCTCGTCTGCTGGTGACCTCATCCCCAACAGTTGGCTCAGGCGATGGCCATCC 230 |
| Db | 341 | CATTTCGTGTCATTCTGGTGACCTCCCACCCTTCAGATGTGAAGGCCAGGCGCATTA 400 |
| QY | 231 | GGCACGCTGGGGAAAAGAGAGGATGGTGAAGGGAAGACGCTGAAGACATTCTTCCTCC 290 |
| Db | 401 | GAGTTACTTGGGTGAAAAAAGTCTTGCTGGGGATATGAGGTTCTTACATTTTCTTAT 460 |
| QY | 291 | TGGGGACCACCGACAGTGCACGCGGAACGAAGAGG-----TGGACCGAGGAGGCC 341 |
| Db | 461 | TAGCCCAAGAGGCTTGAAGAGGAAGACAAATGTTGGCATTTGCTCTAGAGGATGAACAC 520 |
| QY | 342 | AGCGACACGGGGACATTATCCAGAGGATTTCCCTAGACGCTATTACAACTGACCCCTGA 401 |
| Db | 521 | TTCTTTTATGGTGACATAATCCGCAAGATTTTTTAGACACATATATAACCTGACCTTGA 580 |
| QY | 402 | AGACCATATGGGATAGAATGGGTCCATCGCTTTTGTCCCTCAGSGCGGCTTTGTGATGA 461 |
| Db | 581 | AAACCATATGGCATTGAGGTGGGTAACTGAGTTTTGCCCCCAATGCCAAGTACGTAATGA 640 |
| QY | 462 | AACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTTCTTGAAGAAA 521 |

| | | |
|----|--------------|---------------|
| PR | 10-JUN-1998; | 98US-0088740; |
| PR | 10-JUN-1998; | 98US-0088741; |
| PR | 10-JUN-1998; | 98US-0088742; |
| PR | 10-JUN-1998; | 98US-0088810; |
| PR | 10-JUN-1998; | 98US-0088811; |
| PR | 10-JUN-1998; | 98US-0088824; |
| PR | 10-JUN-1998; | 98US-0088825; |
| PR | 10-JUN-1998; | 98US-0088826; |
| PR | 11-JUN-1998; | 98US-0088858; |
| PR | 11-JUN-1998; | 98US-0088861; |
| PR | 11-JUN-1998; | 98US-0088863; |
| PR | 11-JUN-1998; | 98US-0088876; |
| PR | 12-JUN-1998; | 98US-0088876; |
| PR | 12-JUN-1998; | 98US-0089090; |
| PR | 12-JUN-1998; | 98US-0089105; |
| PR | 16-JUN-1998; | 98US-0089440; |
| PR | 16-JUN-1998; | 98US-0089512; |
| PR | 16-JUN-1998; | 98US-0089514; |
| PR | 17-JUN-1998; | 98US-0089532; |
| PR | 17-JUN-1998; | 98US-0089533; |
| PR | 18-JUN-1998; | 98US-0089801; |
| PR | 18-JUN-1998; | 98US-0089907; |
| PR | 18-JUN-1998; | 98US-0089908; |
| PR | 19-JUN-1998; | 98US-0089947; |
| PR | 19-JUN-1998; | 98US-0089948; |
| PR | 19-JUN-1998; | 98US-0089952; |
| PR | 22-JUN-1998; | 98US-0090246; |
| PR | 22-JUN-1998; | 98US-0090252; |
| PR | 23-JUN-1998; | 98US-0090254; |
| PR | 23-JUN-1998; | 98US-0090349; |
| PR | 23-JUN-1998; | 98US-0090355; |
| PR | 24-JUN-1998; | 98US-0090423; |
| PR | 24-JUN-1998; | 98US-0090431; |
| PR | 24-JUN-1998; | 98US-0090435; |
| PR | 24-JUN-1998; | 98US-0090444; |
| PR | 24-JUN-1998; | 98US-0090445; |
| PR | 24-JUN-1998; | 98US-0090472; |
| PR | 24-JUN-1998; | 98US-0090535; |
| PR | 24-JUN-1998; | 98US-0090540; |
| PR | 24-JUN-1998; | 98US-0090557; |
| PR | 25-JUN-1998; | 98US-0090676; |
| PR | 25-JUN-1998; | 98US-0090678; |
| PR | 25-JUN-1998; | 98US-0090688; |
| PR | 25-JUN-1998; | 98US-0090690; |
| PR | 25-JUN-1998; | 98US-0090691; |
| PR | 25-JUN-1998; | 98US-0090694; |
| PR | 25-JUN-1998; | 98US-0090695; |
| PR | 25-JUN-1998; | 98US-0090696; |
| PR | 26-JUN-1998; | 98US-0090862; |
| PR | 26-JUN-1998; | 98US-0090863; |
| PR | 01-JUL-1998; | 98US-0091358; |
| PR | 01-JUL-1998; | 98US-0091360; |
| PR | 01-JUL-1998; | 98US-0091544; |
| PR | 02-JUL-1998; | 98US-0091478; |
| PR | 02-JUL-1998; | 98US-0091486; |
| PR | 02-JUL-1998; | 98US-0091519; |
| PR | 02-JUL-1998; | 98US-0091626; |
| PR | 02-JUL-1998; | 98US-0091628; |
| PR | 02-JUL-1998; | 98US-0091633; |
| PR | 02-JUL-1998; | 98US-0091646; |
| PR | 02-JUL-1998; | 98US-0091673; |
| PR | 07-JUL-1998; | 98US-0091978; |
| PR | 09-JUL-1998; | 98US-0092182; |
| PR | 20-JUL-1998; | 98US-0092472; |
| PR | 20-JUL-1998; | 98US-0093339; |
| PR | 30-AUG-1998; | 98US-0094651; |
| PR | 04-AUG-1998; | 98US-0095282; |

| | | |
|----|--------------|----------------|
| PR | 04-AUG-1998; | 98US-00953288; |
| PR | 04-AUG-1998; | 98US-00953301; |
| PR | 04-AUG-1998; | 98US-00953302; |
| PR | 04-AUG-1998; | 98US-00953318; |
| PR | 04-AUG-1998; | 98US-00953321; |
| PR | 04-AUG-1998; | 98US-00953325; |
| PR | 10-AUG-1998; | 98US-00959916; |
| PR | 10-AUG-1998; | 98US-00959929; |
| PR | 10-AUG-1998; | 98US-00960112; |
| PR | 11-AUG-1998; | 98US-00961413; |
| PR | 11-AUG-1998; | 98US-00961426; |
| PR | 12-AUG-1998; | 98US-00963329; |
| PR | 17-AUG-1998; | 98US-00967571; |
| PR | 17-AUG-1998; | 98US-00967668; |
| PR | 17-AUG-1998; | 98US-00967686; |
| PR | 17-AUG-1998; | 98US-00967773; |
| PR | 17-AUG-1998; | 98US-00967791; |
| PR | 17-AUG-1998; | 98US-00968667; |
| PR | 17-AUG-1998; | 98US-00968691; |
| PR | 17-AUG-1998; | 98US-00968694; |
| PR | 17-AUG-1998; | 98US-00968695; |
| PR | 17-AUG-1998; | 98US-00968697; |
| PR | 18-AUG-1998; | 98US-00969549; |
| PR | 18-AUG-1998; | 98US-00969550; |
| PR | 18-AUG-1998; | 98US-00969559; |
| PR | 18-AUG-1998; | 98US-00969660; |
| PR | 18-AUG-1998; | 98US-00970022; |
| PR | 19-AUG-1998; | 98US-00971481; |
| PR | 20-AUG-1998; | 98US-00972118; |
| PR | 24-AUG-1998; | 98US-00976611; |
| PR | 26-AUG-1998; | 98US-00979511; |
| PR | 26-AUG-1998; | 98US-00979522; |
| PR | 26-AUG-1998; | 98US-00979554; |
| PR | 26-AUG-1998; | 98US-00979571; |
| PR | 26-AUG-1998; | 98US-00979774; |
| PR | 26-AUG-1998; | 98US-00979778; |
| PR | 26-AUG-1998; | 98US-00979798; |
| PR | 26-AUG-1998; | 98US-00979866; |
| PR | 26-AUG-1998; | 98US-00980814; |
| PR | 31-SEP-1998; | 98US-00985825; |
| PR | 16-SEP-1998; | 98US-0100634; |
| PR | 12-JAN-1999; | 99US-0115565; |

(GETH) GENENTECH INC.

Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK; Wood WI, Yuan J;

WPI; 2000-072883/06.

P-PSDB; AAY66685.

Membrane-bound proteins and related nucleotide sequences

Claim 2; Fig 136; 822pp; English.

The invention provides membrane-bound PRO polypeptides and polynucleotides encoding them. The PRO sequences of the invention were identified based on extracellular domain homology screening. The PRO sequences have homology with proteins including LDL receptors, TIE ligands and various enzymes. The membrane-bound proteins and receptor molecules are useful as pharmaceutical and diagnostic agents. Receptor immunoadhesins, for instance, can be used as therapeutic agents to block receptor-ligand interactions. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. The PRO encoding sequences are useful as hybridization probes, in chromosome and gene mapping and in the generation of antisense RNA and DNA. PRO nucleic acid sequences will also be useful for the preparation of PRO polypeptides, especially by recombinant techniques.

Sequence 2095 BP; 629 A; 382 C; 418 G; 666 T; 0 other;

| Query Match | 14.08; | Score 130.4; | DB 21; | Length 2095; |
|-----------------------|---|--------------------|-----------|--------------|
| Best Local Similarity | 48.94; | Pred. No. 1.6e-29; | | |
| Matches 387; | Conservative 0; | Mismatches 396; | Indels 9; | Gaps 1; |
| Qy 111 | TTTAAAGAAAGCGGAAACTTCCTTAAGCTCCAGATACAGACTGCAGGAGAGACACCTC | 170 | | |
| Db 494 | TTTACAGACAGACTTTCACCTTCACACTTCAGAGACATTCAACTGCTCTCATCAAAATC | 553 | | |
| Qy 171 | CCTTCTCGTCTGCTGGTGACCTCATCCCAACAGTTCGGCTGAGCGCATGCCATCC | 230 | | |
| Db 554 | CATTCTGGTCACTCTGGTGACCTCCCAACCTTCAGATGTGAAAGCCAGGCGCACTTA | 613 | | |
| Qy 231 | GGCAGAGCTGGGGAAGAGAGAGATGGTGAAGGAAACAGCTGNAGACATCTTCCTCC | 290 | | |
| Db 614 | GAGTACTTGGGTGAAAAAAGCTTGGTGGGATATGAGGTTCTTACATTTTCTTAT | 673 | | |
| Qy 291 | TGGGACACACAGCAGTGCAGCGGAAAGAAAGG-----TGACACAGAGAGGCC | 341 | | |
| Db 674 | TAGCCAAAGAGGCTGAAAAGGACACAAAATGTTGGCATTCCTTAGAGATGACACC | 733 | | |
| Qy 342 | AGCAGACGGGGACATTCACAGAGGATTTCTTAGAGGCTATTACAACTGACCCCTGA | 401 | | |
| Db 734 | TTCTTTATGGTGACATAATCCGACAAGATTTTATAGACACATATAAATACCTGACCTGA | 793 | | |
| Qy 402 | AGACCATGATGGCATAGAATGGGTCCATCCTTTTGTCCTCAGCGCGCTTTGTGATGA | 461 | | |
| Db 794 | AAACCATATGGCAATTAGGTGGGTACTGAGTTTTGCCCAATGCCAAGTACGATGA | 853 | | |
| Qy 462 | AAACAGACTCAGACATGTTCAATGTGACTATCTGACTGAACTGCTCTGGAAGAAA | 521 | | |
| Db 854 | AGACAGACACTGATGTTTTCATCAATAGTGCATTTAGTCAAGTATCTTTTAAACCTAA | 913 | | |
| Qy 522 | ACAGAACACAGGTTTTTCTACTGGCTCTTGAACCTCAATGAGTTTCCCATCAGGACG | 581 | | |
| Db 914 | ACCACCTCAGAGAAGTTTTTTCACAGGTATCTCTTAATGTGAATTAATTCCTATAGAGGAT | 973 | | |
| Qy 582 | CATTTCAGCAAGTGTGTCTAGTAAATCTGAATCTCCGTGGGACAGGTACCCACCATCT | 641 | | |
| Db 974 | TTTACCAAAAAACCATATTTCTTACCAGGAGTATCCTTTCAAGGTGTCCCTCCATACT | 1033 | | |
| Qy 642 | GCTCCGGCACCGGCTAGTGTTTTCTGGCGAGCTGGCGAGTCAAGTGTACAATGTCTCCA | 701 | | |
| Db 1034 | GCAGTGGGTTGGGTTATATAATGTCCAGAGATTTGGTGCCAAAGATCTATGAAATGATGG | 1093 | | |
| Qy 702 | AGAGCGTCCCATACATTAAACTGGAAGACGTTGTGGGGCTCTGCCTCCAAAGGCTGA | 761 | | |
| Db 1094 | GTACAGTAAACCCATCAAGTTGAAGATGTTTATGCGGGATCTGTTGAATTTATTAA | 1153 | | |
| Qy 762 | ACATCAGATTGGAGGAGCTCCACTCCAGCGGACCTTTTTTCCAGGGGGCTTACGCTTCT | 821 | | |
| Db 1154 | AAGTGAACATTTCATATCCAGAAGACACAAATCTTTTCTTCTATATAGAATCCATTTGG | 1213 | | |
| Qy 822 | CCGTATGCCCTTCCAGGAGGATCGTGCCCTGCCACTTCATCAAGCCCTCGGACTCTCTGG | 881 | | |
| Db 1214 | ATGTCTGTCAACTGAGACGTGTGATTCAGCCCCATGGCTTTCTTCCAAGGAGATCATCA | 1273 | | |
| Qy 882 | ACTACTGGCAGG | 893 | | |
| Db 1274 | CTTTTGGCAGG | 1285 | | |

RESULT 8
AAS46005
ID AAS46005 standard; CDNA; 2095 BP.

DT 18-DEC-2001 (first entry)

DE Human DNA encoding PRO polypeptide sequence #81.

PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep; ss;
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;

| | | |
|----|---------------------------------|--|
| KW | blood; chondrocyte cell; cell | |
| KW | adrenal; lung; breast; prostate | |
| KW | PCR primer. | |
| XX | | |
| OS | Homo sapiens. | |
| XX | | |
| PN | WO200168848-A2. | |
| XX | | |
| PD | 20-SEP-2001. | |
| XX | | |
| PF | 28-FEB-2001; 2001WO-US06520. | |
| XX | | |
| PR | 01-MAR-2000; 2000WO-US05601. | |
| PR | 02-MAR-2000; 2000WO-US05841. | |
| PR | 03-MAR-2000; 2000US-187202P. | |
| PR | 06-MAR-2000; 2000US-186968P. | |
| PR | 14-MAR-2000; 2000US-189320P. | |
| PR | 14-MAR-2000; 2000US-189328P. | |
| PR | 15-MAR-2000; 2000WO-US06884. | |
| PR | 21-MAR-2000; 2000US-190082P. | |
| PR | 21-MAR-2000; 2000US-191007P. | |
| PR | 21-MAR-2000; 2000US-191048P. | |
| PR | 21-MAR-2000; 2000US-191314P. | |
| PR | 28-MAR-2000; 2000US-192655P. | |
| PR | 29-MAR-2000; 2000US-193032P. | |
| PR | 30-MAR-2000; 2000US-193053P. | |
| PR | 30-MAR-2000; 2000WO-US08439. | |
| PR | 04-APR-2000; 2000US-194449P. | |
| PR | 04-APR-2000; 2000US-194647P. | |
| PR | 11-APR-2000; 2000US-195975P. | |
| PR | 11-APR-2000; 2000US-196000P. | |
| PR | 11-APR-2000; 2000US-196187P. | |
| PR | 11-APR-2000; 2000US-196690P. | |
| PR | 11-APR-2000; 2000US-196820P. | |
| PR | 18-APR-2000; 2000US-198121P. | |
| PR | 18-APR-2000; 2000US-198585P. | |
| PR | 25-APR-2000; 2000US-199397P. | |
| PR | 25-APR-2000; 2000US-199550P. | |
| PR | 25-APR-2000; 2000US-199654P. | |
| PR | 03-MAY-2000; 2000US-201516P. | |
| PR | 17-MAY-2000; 2000WO-US13705. | |
| PR | 22-MAY-2000; 2000WO-US14042. | |
| PR | 30-MAY-2000; 2000WO-US14941. | |
| PR | 02-JUN-2000; 2000WO-US15264. | |
| PR | 05-JUN-2000; 2000US-203832P. | |
| PR | 28-JUL-2000; 2000WO-US20710. | |
| PR | 22-AUG-2000; 2000US-0644848. | |
| PR | 24-AUG-2000; 2000WO-US23328. | |
| PR | 08-NOV-2000; 2000WO-US30952. | |
| PR | 01-DEC-2000; 2000WO-US32678. | |
| PR | 20-DEC-2000; 2000WO-US34956. | |

(GETH) GENENTECH INC.

Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL; Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;

WPI; 2001-602746/68.
P-PSDB: AAU29104.

Novel nucleic acids encoding PRO polypeptides, used to diagnose the presence of tumours, such as prostate and breast tumours, in mammals and to screen for modulators of the compounds -

Claim 2; Fig 161; 774pp; English.

Sequences AAS45925-AAS46231 represent DNA molecules encoding and PCR primers for PRO polypeptides of the invention. The sequences of the invention can be used to detect the presence of a tumour in a mammal by comparing the level of expression of a PRO polypeptide in a test sample of cells from the animal and a control sample of normal cells, whereby a higher level of expression in the test sample indicates the presence of a tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep,

CC pigs, goats and rabbits but are preferably human. The polypeptides can be
CC used to stimulate tumour necrosis factor (TNF) alpha release from human
CC blood, when contacted with it. A specific polypeptide can be used to
CC stimulate the proliferation or differentiation of chondrocyte cells. The
CC PRO proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adrenal, lung, colon,
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders.
XX
XX Sequence 2095 BP: 629 A; 382 C; 418 G; 666 T; 0 other;

SQ Sequence 2095 BP; 629 A; 382 C; 418 G; 666 T; 0 other;

| | | | | |
|-----------------------|--------------|--------------------|-----------------|-------------------|
| Query Match | 14.0%; | Score 130.4; | DB 22; | Length 2095; |
| Best Local Similarity | 48.9%; | Pred. No. 1.6e-29; | | |
| Matches 387; | Conservative | 0; | Mismatches 396; | Indels 9; Gaps 1; |

| | | | |
|----|------|--|------|
| Qy | 111 | TTTACAGAAAGACGGAACTTCTCTTAAGCTCCCGAGATACAGACTGCAGGAGACACCTC | 170 |
| | | | |
| Db | 494 | TTTACAGCAAGAACTTTCACTTCACACTTCGAGAGCACTTCAAACTGCCTCTCATCAAAAT | 553 |
| | | | |
| Qy | 171 | CCTTCTCGTCCTGCTGGTGACCTCATCCCAAAACAGTTGGCTCAGCGCATGGCCATCC | 230 |
| | | | |
| Db | 554 | CATTTCTGGTCATCTTGGTGACCTCCCACTTCAGATGTTGAAAGCCAGGAGGCCATTA | 613 |
| | | | |
| Qy | 231 | GGCAGACGTGGGGGAAAGAGAGATGGTGAAGGGAAGCAGCTGAAGACATCTTCTCCCTCC | 290 |
| | | | |
| Db | 614 | GAGTTACTTGGGTGAAAAAAAGTCTTGCTGGGGATATGAGGTTCTTACATTTTTCTTAT | 673 |
| | | | |
| Qy | 291 | TGGGGACCAACCACTGCAGCGGAACAAAGAGG-----TGGACCAAGGAGACC | 341 |
| | | | |
| Db | 674 | TAGGCCAAGAGGCTGAAAAGGAAGCAAAATCTTGGCAATGTGCTTAGAGGATGAACACC | 733 |
| | | | |
| Qy | 342 | AGGCACAGGGGACATATTCCAGAGGATTTCTTAGACGCTCATTTACAATCTGACCCCTGA | 401 |
| | | | |
| Db | 734 | TTCTTTTGGTGACATAATCCGACAGAATTTTTAGACACATATAATTAACCTGACCTTGA | 793 |
| | | | |
| Qy | 402 | AGACCATGTGGGCATAGAAATGGTCCATCGCTTTTGTGCTCAGCGGCGGTTTGTGATGA | 461 |
| | | | |
| Db | 794 | AAACCAATATTGGCATTCAGGTGGGTAACTGAGTTTGTGCCCAATGCCAAGTACGTAATGA | 853 |
| | | | |
| Qy | 462 | AAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTCAACTGCTTCTCGAAGAAA | 521 |
| | | | |
| Db | 854 | AGACAGACACTGATGTTTTTCATCACTACTGGCAATTTAGTGAAGTATCTTTTTAAACCTAA | 913 |
| | | | |
| Qy | 522 | ACAGAAACAACAGGTTTTTCACTGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGCAGC | 581 |
| | | | |
| Db | 914 | ACCACTCAGAGAAGTTTTTTCACAGGTTATCCTCTAATTGATAATTATTCCTATAGAGAT | 973 |
| | | | |
| Qy | 582 | CATTACGACAGTGGTTTGTTCAGTAATCTGAATATCCGTGGGACAGGTACCCACCATCT | 641 |
| | | | |
| Db | 974 | TTTACCAAAAACCCCATATTTCTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCCATCT | 1033 |
| | | | |
| Qy | 642 | GCCTCCGCGCACGGCTACGTGTTTTCTTGGCGCGTGGCGAGTCAGGTGTACAATGCTCCA | 701 |
| | | | |
| Db | 1034 | GCACTGGGTTGGGTTATATAATGTCAGAGATTTGGTGCCANGGATCTATGAATGATGG | 1093 |
| | | | |
| Qy | 702 | AGAGCTGCCATACATTAACCTGGAAGAGCGTGTTTGTGGGCTCTGCCCTCGAAAGGCTGA | 761 |
| | | | |
| Db | 1094 | GTCACTGTAACCCCAAGTTTGAAGATGTTTATGTCCGGATCTGTTTGAATTTATTA | 1153 |
| | | | |
| Qy | 762 | ACATCAGATTGGAGGACTCCACTCCCGCCGACCTTTTTTCCAGGGGGCTTACGCTTCT | 821 |
| | | | |
| Db | 1154 | AAGTGAACATTCATATPCCAGAAGACACAAATCTTTTTCTTCTATATAGAATCCATTTGG | 1213 |
| | | | |
| Qy | 822 | CCGTATGCCCTCTTCAGGAGGATCGTGGCTCGCACTTCATCAAGGCTCGGACTCTCTTGG | 881 |
| | | | |
| Db | 1214 | ATGCTGTCACTGAGACGTGTGATTGACGCCCATGGCTTTTCTTCCAAGGAGATCATCA | 1273 |
| | | | |
| Qy | 882 | ACTACTGGCAGG | 893 |
| | | | |
| Db | 1274 | CTTTTGGCAGG | 1285 |

RESULT 9
AAAF92075
ID AAAF92075 standard; cDNA; 2095 BP.

AA AAF92075;

XX
DT 15-MAY-2001 (first entry)XX
DE
Human PR01074 cDNA

XX
KW Human: PRO protein: manning: ss

XX
50
Homosapiens

XX
DN W0200116319-82

XX
00-432-2001

XX

XX

PR 15-SEP-1999; 99WO-US21090.

PR 09-DEC-1999; 99US-0170262.

PR 18-FEB-2000; 2000WO-US04341.

PR 2000WO-US04414;
22-FEB-2000;
19 FEB 2000; 2000WO-US04342;
PR

01-MAR-2000; 2000WC-0303801.
03-MAR-2000; 2000US-0187202.

PR 23-APR-2000; 2000US-0199397.
PR 22-MAY-2000; 2000WO-US14042.

PR 05-JUN-2000; 200005-0209832.
XXPA (GETH) GENENTECH INC.
XX

PI Eaton BL, Filvaroff E, Ger
PI Crimaldi CI, Curran M, Wa

XX
 WPT: 2001-102260/10

DR P-PSDB; AAB87543.

PT Eighty four nucleic ac

PT chromosome and gene mapping. -

PS Claim 2; Fig 35; 278pp; English.

CC The present sequence

CC antagonists or an

CC measurement useful in the treatment of a condition which is responsive to
CC the PRO protein, agonists, antagonists or anti-PRO antibodies. The PRO
CC protein may also be employed as molecular weight markers for protein
CC electrophoresis. The PRO coding sequence has applications in molecular
CC biology, including use as hybridisation probes, and in chromosome and
CC gene mapping.

Sequence 2095 BP: 629 A: 382 C: 418 G: 666 T: 0 other:

Query Match 14.0%; Score 130.4; DB 22; Length 2095;
Best Local Similarity 48.9%; Pred. No. 1.6e-29;
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps

QY 111 TTTCAGAAAGACGGGAACCTTCCTTAAGCTCCAGATACAGACTGCAGGCAGACACCTC 170

| Accession | Gene | Strain | Sequence | Position |
|-----------|------|--------|--|----------|
| 494 | Db | 494 | TTTACAGACAAGACTTTCACCTTCGAGAGCATTTCAACATGCTCATCAAAATC | 553 |
| 171 | QY | 171 | CCTTCCTCGTCGCTGGTGACCTCATCCCAAAACAGTTGGCTGAGCGCATGGGCATCC | 230 |
| 554 | Db | 554 | CATTTCTGGTCATCTTGGTGACCTCCACCCCTTCAGATGAAAGCCAGGCAGGCCATTA | 613 |
| 231 | QY | 231 | GGCAGACGTGGGGGAAAGAGAGGATGGTGAAGGGAAGCAGCTGAAGACATTTCTTCCTCC | 290 |

Db 614 GAGTTACTTGGGTCAGAAAGAGTCTTGGTGGGATATCAGGCTTCTTACATTTTCTTAT 673
Qy 291 TGGGGACACACAGAGTGCAGCGGAAAGAGG-----TGGACAGGAGGCC 341
Db 674 TAGGCCAAGAGGCTGAAAGGAAGACAAAATGTTGGCATTGTTCCTTAGAGGATGAACACC 733
Qy 342 AGCGACAGGGGACATATCCAGAGGATTTCTAGAGCTCTATTACATCTGACCCCTGA 401
Db 734 TTCTTTATGGTGACATATCCGACAGATTTTTAGACACATATAATACCTGACCTTGA 793
Qy 402 AGACCATGATGGGCATAGAAATGGGTCATCGCTTTTGTCTCAGCGGGCTTTGTCATGA 461
Db 794 AAACCATATGACATTCAGGTGGTAACTGAGCTTTTGGCCCAATGCCAAGTACGTAATGA 853
Qy 462 AAACAGACATCAGACATGTTTCATCATGTTGAGTATCTGACTGAACTGCTTCTGAAGAAA 521
Db 854 AGACAGACACTGATGTTTTCATCAATGTTGAGTATCTGACTGAACTGCTTCTTAAACCTAA 913
Qy 522 ACAGAACCAACAGGTTTTTCTACTGGCTTCTTGAAGTCTCAATGAGTTTCCCATCAGGCAGC 581
Db 914 ACCACTCAGAGAGTTTTCACAGGTTATCTCTTAATGATATATTTCTCTATAGAGAT 973
Qy 582 CATTCAGCAAGTGGTTTGTGCTAATATCTGAATATCCGTGGGACAGGTACCCACCATCT 641
Db 974 TTTACCAAAAACCCATATTTCTTACCAGGAGTATCTTTCAAGGTGTTCCTCCATACT 1033
Qy 642 GCTCGGACCGCTACGTTTCTTCTGCGAGCTGCGAGTGGGATGTTGAGTATGTTCTTCA 701
Db 1034 GCAGTGGGTGGTATATAATGTCAGAGATTTGGTCCAGGATCTATGAAATGATGG 1093
Qy 702 AGAGGCTCCCATACATTAACCTGGAAGAGCTGTTTGTGGGCTCTGCCTCGAAAGGCTGA 761
Db 1094 GTCACGTAACCCATCAAGTTTGAAGATGTTTATGTCGGGATCTGTTTGAATTTATTA 1153
Qy 762 ACATCAGATTGGAGAGCTCCACTCCAGCCGACCTTTTTCAGGGGCTTACGCTTCT 821
Db 1154 AAGTGAACATTCATATCCAGAGACACAAAATCTTTTCTCTATATAGAAATCCATTTGG 1213
Qy 822 CCGTATGCTCTTCAGGAGGATCGTGGCTGCACTTTCATCAAGCTCGGACTCTCTTGG 881
Db 1214 ATGCTGTCACAGTGAAGCTGATTTGACGCGCCATGGCTTTCTTCCAGGAGATCATCA 1273
Qy 882 ACTACTGGCAGG 893
Db 1274 CTTTTTGGCAGG 1285

RESULT 10

AAF44168 ID AAF44168 standard; cDNA; 2095 BP.

XX AC AAF44168;

XX DT 02-APR-2001 (first entry)

XX DE Human PRO1074 (UNQ531) nucleotide sequence SEQ ID NO:208.

XX KW Human; secreted and transmembrane protein; PRO; cytosstatic;

XX KW cell death; cancer; chromosomal mapping; gene mapping; tissue typing;

XX KW diagnostic assay; ss.

XX OS Homo sapiens.

XX PN WO200073454-A1.

XX PD 07-DEC-2000.

XX PF 30-MAR-2000; 2000WO-US08439.

XX PR 02-JUN-1999; 99WO-US12252.

XX PR 23-JUN-1999; 99US-0141037.

XX PR 07-JUL-1999; 99US-0143048.

PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 28-JUL-1999; 99US-0148222.
PR 17-AUG-1999; 99US-0149396.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 08-OCT-1999; 99WO-US158663.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 20-MAR-2000; 2000WO-US07377.
XX (GETH) GENENTECH INC.
PA Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi CJ, Gurney AL, Kijavlin IJ, Napier MA, Pan J, Paoni NF;
PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;
PI Zhang Z;
XX WPI; 2001-032160/04.
DR P-PSDB; AAB65208.

XX PRO polynucleotides used to produce polypeptides used to target
PT bioactive molecules such as toxins, radiolabels or antibodies, to
PT specific cells, to cause targeted cell death -

PS Claim 2; Fig 136; 935pp; English.

XX The present invention describes human secreted and transmembrane PRO
CC proteins. The PRO proteins have cytostatic activity. The PRO proteins
CC can be used for targeted delivery of bioactive molecules, such as
CC toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide
CC sequences, and their fragments, can be used as hybridisation probes, in
CC chromosomal and gene mapping, and in the generation of anti-sense RNA
CC and DNA. They may also be used to produce transgenic animals which are
CC used to develop and screen therapeutically useful reagents. The PRO
CC nucleotide and protein sequence can be used for tissue typing and in
CC treating cancer. Anti-PRO antibodies can be used in diagnostic assays.
CC AAF44270 to AAF44470 represent PCR primers and hybridisation probes used
CC in the isolation of human PRO sequences. AAF44087 to AAF44269 and
CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein
CC sequences given in the exemplification of the present invention.
XX Sequence 2095 BP; 629 A; 382 C; 418 G; 666 T; 0 other;

Qy 111 TTTACAGAAAGACGGGAACTTCTTAAAGTCCCGATACAGACTGACGACGACACCTC 170
Db 494 TTTACAGACAAGACTTTTCACTTCAGAGGATTCACAACTGCTCTCATCAAAATC 553
Qy 171 CGTTCCTCGTCTGCTGAGCTCATCCACAAAGTGGCTGAGCGCATGCCATCC 230
Db 554 CATTTCTGCTCATCTTGGTGGCTCCACCTTCAGATGTGAAGCCAGGAGCCATTA 613
Qy 231 GGCAGACGTGGGGGAAAGAGAGATGGTGAAGGAAAGCAGCTGAAGACATCTTCTCTCC 290
Db 614 GAGTTACTTGGGGTGAAAAAAGTCTTGGTGGGATATGAGGTTCTTACATTTTCTTAT 673
Qy 291 TGGGGACCAACAGCAGTGCAGCGGAAAGAGG-----TGGACAGGAGGCC 341

Query Match 14.0%; Score 130.4; DB 22; Length 2095;

Best Local Similarity 48.9%; Pred. No. 1.6e-29;

Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;

Db 674 TAGGCCAAGAGCGCTGAAAGGAAGACAAAATCTTGCCATTTGCTTTAGAGGATGAACACC 733
Qy 342 ACGGACACGGGACACATTATCCAGAGAGGATTTCTTACAGCTCTATTACAATCTGACCCCTGA 401
Db 734 TTCTTTATGGTGACATAATCCGACAGATTTTATAGACACATATAATAACCTGACCTTGA 793
Qy 402 AGACCATGATGGCATAGAAATGGGTCCTATCGCTTTTGTCTCAGGCGGGCTTTGTGATCA 461
Db 794 AAACCATATGCAATTCAGGTGGGTAACTGAGTTTGCCCAATGCCAAGTACGTAATGA 853
Qy 462 AAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACGTCTTCTCAAGAAAA 521
Db 854 AGACAGACACTGATGTTTTCATCAATACTGCGCAATTTAGTGAAGTATCTTTTAAACCTAA 913
Qy 522 ACAGAACACACAGGTTTTCACCTGGCTTCTGAAACTCAATGAGTTTCCCATCAGGCAGC 581
Db 914 ACCACTCAGAGAGGTTTTTTCACAGGTTATCTCTTAATGATAAATATTCCCTATAGAGAT 973
Qy 582 CATTCAGCAAGTGGTTTCTCAGTAATCTGAATATCCGTGGGACAGGTACCCACCATTTCT 641
Db 974 TTTACCAAAAACCCCATATTTCTTACCAGGACTATCTTTCAAGGTGTTCCCTCCATATCT 1033
Qy 642 GCTCCGGACCGGCTACGTTTCTTGGGAGCGTGGCGAGTCAGGTGTAACAATGTCTCA 701
Db 1034 GCAGTGGGTTGGTTATATAATCTCCAGAGATTTGGTCCCAAGGATCTATGAATGATGG 1093
Qy 702 AGAGCGTCCATACATTAACCTGGAAGAGCGTGTGTTGGGGCTCTGCCCTCGAAGGCTGA 761
Db 1094 GTCAGCTAAACCCCATCAAGTTTGAAGATGTTTATGTCGGGATCTGTTTGAATTTATTA 1153
Qy 762 ACATCAGATTTGGAGGAGCTCCACTCCAGCCGACCTTTTTCAGGGGCTTTACGCTTCT 821
Db 1154 AAGTGAACATTCATATCCAGAGACACAAAATCTTTCTTCTATATAGAATCCATTTGG 1213
Qy 822 CGGTATGCTCTTCAGGAGGATCGTGGCTGCCACTTCATCAAGCTCGGACTCTCTTTGG 881
Db 1214 ATGCTGTCACTGAGAGCGTGTGATTGGAGCCCATGGCTTTCTTCCCAAGGAGATCATCA 1273
Qy 882 ACTACTGGCAGG 893
Db 1274 CTTTGTGGCAGG 1285

RESULT 11
AAH15711
ID AAH15711 standard; cDNA: 2168 BP.
XX
AC AAH15711;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:14092.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isodai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI: 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs.
XX
PS Claim 8: SEQ ID 14092; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
full-length cDNAs defined in the specification. Where a primer set
comprises: (a) an oligo-dT primer and an oligonucleotide complementary
to the complementary strand of a polynucleotide which comprises one of
the 5602 nucleotide sequences defined in the specification, where the
oligonucleotide comprises at least 15 nucleotides; or (b) a combination
of an oligonucleotide comprising a sequence complementary to the
complementary strand of a polynucleotide which comprises a 5'-end
sequence and an oligonucleotide comprising a sequence complementary to a
polynucleotide which comprises a 3'-end sequence, where the
oligonucleotide comprises at least 15 nucleotides, and the combination of
the 5'-end sequence/3'-end sequence is selected from those defined in
the specification. The primer sets can be used in antisense therapy and
in gene therapy. The primers are useful for synthesizing polynucleotides,
particularly full-length cDNAs. The primers are also useful for the
detection and/or diagnosis of the abnormality of the proteins encoded by
the full-length cDNAs. The primers allow obtaining of the full-length
cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
represent oligonucleotides, all of which are used in the exemplification
of the present invention.
SQ Sequence 2168 BP; 634 A; 405 C; 452 G; 677 T; 0 other;
Query Match 14.0%; Score 130.4; DB 22; Length 2168;
Best Local Similarity 48.9%; Pred. No. 1.7e-29;
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;
Qy 111 TTTACAGAAAGACGGGAACCTTCTTAAGTCCCGATACAGACTGCAGCAGACACCTC 170
Db 571 TTTACAGACAAGACTTTTCACTTCACACTTCGAGAGCATTCAAACTGCTCTCATCAAAATC 630
Qy 171 CCTTCTCTCTCTGCTGACCTCATCCACAAACACTTGGCTGAGCGGATGGCATCC 230
Db 631 CATTTCTGGTCTATCTGCTGACCTCCACCCCTTCAGATGTGAAGGCCAGCAGCGCATTA 690
Qy 231 GGCAGACGTGGGGGAAAGAGAGGATGGTGAAGGGAAGAGAGCTGAAGACATCTTCTCTCC 290
Db 691 GAGTTACTTGGGTGAAAAAAGTCTTGGTGGGATATAGGTTCTTACATTTTCTTAT 750
Qy 291 TGGGGACCCAGCAGCTGCAGCGGAACGAAAGAGG-----TGGACAGGAGAGCC 341
Db 751 TAGGCCAAGAGCTGAAAAGGAAGACAAAATGTTGCCATTGCTCTTAGAGGATGAACACC 810
Qy 342 ACCGACACGGGACATATTCCAGAGAGGATTTCTTACAGCTCTATTACAATCTGACCCCTGA 401
Db 811 TTCTTTATGGTGACATAAATCCGACAGATTTTGTAGACACATATAATAACCTGACCTTGA 870
Qy 402 AGACCATGATGGCATAGAAATGGGTCCTATCGCTTTTGTCTCAGGCGGGCTTTGTGATGA 461
Db 871 AAACCATATGCAATTCAGGTGGGTAACTGAGTTTGCCCAACGCAAGTACGTAATGA 930
Qy 462 AAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACGTCTTCTCAAGAAAA 521
Db 931 AGACAGACACTGATGTTTTCATCAATACTGCGCAATTTAGTGAAGTATCTTTTAAACCTAA 990
Qy 522 ACAGAACACACAGGTTTTCACCTGGCTTCTGAAACTCAATGAGTTTCCCATCAGGCAGC 581
Db 991 ACCACTCAGAGAGGTTTTTTCACAGGTTATCTCTTAATGATAAATATTCCCTATAGAGAT 1050
Qy 582 CATTCAGCAAGTGGTTTCTCAGTAATCTGAATATCCGTGGGACAGGTACCCACCATTTCT 641

Db 1051 TTTACCAAAAACCCATATTTCTTACCAGGAGTATCTCTTTCAAGGTGTCCCTCCATACT 1110
Qy 642 GCTCCGGCACCGCTACGTGTTTCTGGCGACGTGGCGAGTCAGTGTCACAAATGTCCTCCA 701
Db 1111 GCATGGGTGGTATATATATGTCAGAGATTTGGTGCCAAAGATCTATGAATGATGG 1170
Qy 702 AGACGGTCCCATACATTAATACTGAAGACGCTGTTGGGGCTCTGCTCGAAAGGCTGA 761
Db 1171 GTCACGTAAAACCCATCAAGTTTCAAGATGTTTATGTCGGGATCTGTTTGAATTTATTAA 1230
Qy 762 ACATCAGATGGAGAGCTCCACTCCAGCGACCTTTTTCAGGGGGCTTACGCTTCT 821
Db 1231 AAGTGAACATTCATATTCAGAGACACAAATCTTTCTATATAGAAATCCATTGG 1290
Qy 822 CCGTATGCCCTCTTCAGGAGGATCGTGGCTGCCACTTCATCAAGCCCTCGGACTCTCTGG 881
Db 1291 ATGCTGTCAACTGAGAGCTGTGATTGCAGCCCATGCTTTCTTCCAAAGGAGATCATCA 1350
Qy 882 ACTACTGGCAGG 893
Db 1351 CTTTTGGCAGG 1362

RESULT 12

AX87192
ID AX87192 standard; cDNA; 2189 BP.
XX
AC AX87192;
XX
DT 27-SEP-1999 (first entry)
XX
DE Human Dendriac cDNA.
XX

KW Dendriac; Brainlac-2; human; Notch; immune disorder;
KW neurological disorder; diagnosis; therapy; ss.
XX
OS Homo sapiens.

Key Location/Qualifiers
FT CDS 435..1394
FT /tag= a
FT sig_peptide 435..509
FT /tag= b
FT mat_peptide 510..1391
FT /tag= c
XX

PN W09931116-AL.

XX 24-JUN-1999.

XX 17-DEC-1998; 98WO-US27049.

XX 17-NOV-1998; 98US-0108928.

XX 18-DEC-1997; 97US-0068006.

XX 12-MAR-1998; 98US-0077687.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ebner R, Endress GA, Florence KA, Rosen CA, Ruben SM;
PI Soppet DR, Yu G;

XX WPI; 1999-430031/36.
DR P-PSDB; AAY06461.

XX Human proteins for treating and detecting immune and neurological
PT disorders

XX Claim 2; Page 110-112; 132pp; English.

XX This is the nucleotide sequence of cDNA clone HFVIF40, which codes
CC for human Dendriac (see AAY06461), also called Brainlac-2, a novel
CC member of the Brainiac family. Dendriac cDNA is deposited as

CC ATCC 203056 and ATCC 209627. The cDNA was initially discovered in
CC a cDNA library derived from dendritic cells. Additional clones of
CC the same gene were identified in NTERA2 cell, adult pulmonary
CC tissue, salivary gland, ovary, Caco-2 colon adenocarcinoma, smooth
CC muscle, cerebellum, 8-week-old human midyo, haemaglobinoma,
CC amygdala, substantia nigra and whole brain cDNA libraries. Dendriac
CC nucleic acids are used in the recombinant production of Dendriac
CC polypeptides, especially the mature protein or epitope-bearing
CC fragments. Dendriac and Brainiac proteins are members of the Notch
CC family of proteins involved in the control of cell division. The
CC proteins may be used to detect and/or treat immune system and
CC neurological disorders.
XX
SQ Sequence 2189 BP; 657 A; 402 C; 453 G; 677 T; 0 other;

Query Match 14.0%; Score 130.4; DB 20; Length 2189;
Best Local Similarity 48.9%; Pred. No. 1.7e-29;
Matches 38; Conservative 0; Mismatches 396; Indels 9; Gaps 1;
Qy 111 TTTACAGAAAGACGGGAACCTTCTTAAGCTCCCAAGATACAGACTGCAGGAGACACCTC 170
Db 571 TTTACAGACAAAGACTTTTCACCTTCAGACTTCGAGAGCAATCAAACTGCTCTCATCAATC 630
Qy 171 CTTTCTCTCTCTGCTGCTGACCTCATCCACAAACAGTTGGCTGAGCGCATGGCCATCC 230
Db 631 CATTTCTGGTCAATCTTGGTGACCTCCACCTTCAGATGTGAAAGCCAGGCGCATTA 690
Qy 231 GCGAGAGTGGGGAAGAGAGAGGATGGTGAAGGGAACAGCTGAAGACATTTCTTCTCTCC 290
Db 691 GAGTTACTTGGGGTGAAAAAAGTCTTGGTGGGATATGAGGTTCTTACATTTTCTTAT 750
Qy 291 TGGGGACACCAGCAGTGCAGCGGAACGAAAGAGG-----TGGACCAGGAGAGCC 341
Db 751 TAGGCAAGAGGCTGAAAGGAAGACAAATGTTGGCATTTGTCCTTAGAGGATGAACACC 810
Qy 342 AGCGACACGGGACATTTATCCAGAGGATTTCTTAGAGCTATTTACAACTTGACCCCTGA 401
Db 811 TTCCTTTATGGTGACATAATCCGACAAGATTTTTCAGACATATAATACCTGACCTTGA 870
Qy 402 AGACCATGATGGGATAGAATGGGTCCATCGCTTTTGGCTCAGCGGGGTGTGTGATGA 461
Db 871 AAACCATATGGCATTCAGGTGGGTAACTGAGTTTGGCCCAATGCCAAGTACGTAATGA 930
Qy 462 AAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACCTGCTTCTGAAGAAA 521
Db 931 AGACAGACACTGATCTTTTCATCAATCTGGAATTTAGTGAAGTATCTTTTAAACCTAA 990
Qy 522 ACAGAACACCCAGCTTTTTCACCTGCTCTTGAACCTCAATGAGTTTCCCATCAGGCAGC 581
Db 991 ACCACTCAGAGAAGTTTTCACAGGTTATCTCTAAATTTGATTAATTTCTTATAGAGAT 1050
Qy 582 CATTCAGCAAGTGGTGTGTCTAGTAAATCTGAAATATCCGTGGGACAGTACCACCATTTCT 641
Db 1051 TTTACCAAAAACCCATATTTCTTACCAGGAGTATCTCTTCAAGGTGTTCCTCCCTCACTACT 1110
Qy 642 GCTCCGGCACCGGCTACGTGTTTCTGGCGACGTGGCGAGTCAGGTGTACAAATGTCCTCA 701
Db 1111 GCATGGGTGGGTATATATAATGTCAGAGATTTTGGTGCCAAAGATCTATGAAATGATGG 1170
Qy 702 AGACGGTCCCATACATTAATACTGAAGACGTTGTTTGGGGCTCTGCTCCCAAGGGCTGA 761
Db 1171 GTCAGGTAAAACCCATCAAGTTTGAAGATGTTTATGTCGGGATCTGTTTGAATTTATTAA 1230
Qy 762 ACATCAGATTTGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCT 821
Db 1231 AAGTGAACATTCATATTTCCAGAGACACAAATCTTTCTTCTATATAGAAATCCATTGG 1290
Qy 822 CCGTATGCCCTCTTCAGGAGGATCGTGGCTGCCACTTCATCAAGCCCTCGGACTCTCTGG 881
Db 1291 ATGCTGTCAACTGAGAGCTGTGATTGCAGCCCATGCTTTCTTCCAAAGGAGATCATCA 1350
Qy 882 ACTACTGGCAGG 893
| | | | |

Db 1351 CTTTTGGCAGG 1362

RESULT 13

AA35710

ID AAX35710 standard; cDNA: 1266 BP.

XX AC AAX35710;

XX DT 09-JUL-1999 (first entry)

XX DE cDNA encoding a protein identified by the signal sequence trap method.

XX KW Signal sequence trap method; SST method; immunisation; inhibition;
 KW infection; allergy; cancer; regulation; tissue formation; tissue repair;
 KW activin activity; inhibin activity; chemokine activity;
 KW cytokine activity; blood coagulation regulation; agonist; antagonist;
 KW metabolic disorder; hormonal disorder; immune disorder;
 KW severe combined immunodeficiency; SCID; AIDS; thrombosis; cancer;
 XX wound; ss.

XX OS Homo sapiens.

XX PN W09918126-A1.

XX PD 15-APR-1999.

XX PF 06-OCT-1998; 98WO-JP04514.

XX PR 07-OCT-1997; 97JP-0274674.

XX PA (ONOY) ONO PHARM CO LTD.

XX PI Fukushima D, Shibayama S, Tada H;

XX DR WPI; 1999-277254/23.

XX DX P-PSDB; AAY02366.

XX PT Polypeptides identified by the signal sequence trap method from a
 human cDNA library

XX PS Claim 4; Page 108-109; 281pp; Japanese.

XX CC AAX35694-X35747 represent cDNA sequences that encode novel polypeptides
 CC (AAY02358-84) which are identified from a human placental cDNA library
 CC by the signal sequence trap (SST) method. The polypeptides have a
 CC broad range of physiological activity, including immunisation against
 CC and inhibition of infections, allergies and cancer; regulation of tissue
 CC formation and repair; activin/inhibin activity; chemokine/cytokine
 CC activity; blood coagulation regulation; and receptor/ligand agonist
 CC or antagonist activity. The polypeptides can be used for prevention
 CC and treatment of disorders including infections by bacteria, yeasts and
 CC viruses (including HIV) and protozoa; metabolic and hormonal disorders;
 CC immune disorders (including severe combined immunodeficiency (SCID)
 CC and AIDS; thrombosis; cancer; and traumatic or surgical wounds.

XX SQ Sequence 1266 BP; 408 A; 270 C; 242 G; 346 T; 0 other;

Query Match 13.18; Score 122.4; DB 20; Length 1266;
 Best Local Similarity 50.48; Pred. No. 3.7e-27;
 Matches 384; Conservative 0; Mismatches 366; Indels 12; Gaps 3;

QY 151 AGACTGAGGACACACCTCCCTCCTGCTGCTGACCTCATCCCAACACAGTT 210

Db 432 AAAATGCCAAGAGAAAGTCTCTTTTAACTACTACTATAGCTGCAGAGCTGGCAAAAT 491

QY 211 GCCTGAGCGCATGCCATCCGGCAGACGTGGGGAAAGAGAGGATGGTGAAGGGAAGCA 270

Db 492 AGAAGCTAGAAGAGCTATTTCGGCAAACTTGGGGCAATGAAAGTCTAGCACCTGGTATTCA 551

QY 271 GCTGAAGACATTTCTCTCTCTGGGACCAACA-GCAGTGCAGCGGAACGAAGAGGTGG 329

Db 552 AATCACAAGATATTTTGTGGCTTAAAGTATTAAAGCTAAATGGCTACCTTCAACGTGC 611

QY 330 A-----CCAGGAGAGCCAGCGACACGGGACATTATCCAGAGGATTTCTCTAGACGCTA 384
 Db 612 AATACTGGAGAAAGAGCAGCAATATCATGATATAATTCACAGGAATACTTAGTAGCTA 671
 QY 385 TTACAATCTGACCCTGAGACCACTGATGGGCATAGAGTGGTCCATCGCTTTTGTCTCTCA 444
 Db 672 CTATATTTTGACCACTTAAACACTAATGGCATGAATGGGTGCAACACTACTGCTCCACA 731
 QY 445 GCGGGCGTTTGTGATGAAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGA 504
 Db 732 TATTCATATGTTATGAAAACTGACAGTGACATGTTTGTCAACACTGAATATTTAATCAA 791
 QY 505 ACTGCTCTCTGAAGAAAACAGAACACCAG-----GTTTTTCACCTGGCTTCTTGAAC 558
 Db 792 TAAGTTACTGAAGCCAGATCTGCCTCCAGACATAACTATTTTCTGCTGTTTACCTAATGCG 851
 QY 559 CAATGAGTTTCCCATCAGGCAGCAATTCAGCAAGTGGTTTGTGAGTAAATCTGAATATCC 618
 Db 852 AGGATATGACCCCAATCGAACAAGATAGCAAGTGGTACATGCCACCAGACCTCTACCC 911
 QY 619 GTGGGACAGGTACCCACCAATCTCTGCTCCGGCACCGGTCTTTTCTGGCGACGTGGC 678
 Db 912 AAGTGAGCGTTATCTCTCTCTGTTCTGGAAGTGGTATGTTTCTGAGATCTGGC 971
 QY 679 GAGTCAGGTGTACAATGTTCTCCAGAGCGTCCCATACATTAACACTGGAAGACGTTTGT 738
 Db 972 AGAAAAGATTTTAAAGTTTCTTAGTATCCGCCCTTTGCACTTGAAGATGTATATGT 1031
 QY 739 GGGGCTCTGCTCGAAAGCTGAACATCAGATTGGAGGAGCTCCACTCCCGACGACCTT 798
 Db 1032 AGGATCTGCTTGGCAAGTTGAGAATGATCTGTACCCCTCCCAATGAGTTTGTGT 1091
 QY 799 TTTTCCAGGGGCTTACGCTTCTCCGTATGCTCTTTCAGGAGGATGCTGGCCTGCCACTT 858
 Db 1092 CAATCACTGGCGAGTCTCTTATTTCGAGCTGTAATACAGCCACTAATTACCTCTCATCA 1151
 QY 859 CATCAAGCCTCGGACTCTCTTGGACTACTGGCAGGCTCTAGA 900
 Db 1152 GTTCCAGCGCTAGTGAACACTGATAAAATACTGGAACCACTTTACA 1193

RESULT 14

AA35711

ID AAX35711 standard; cDNA: 2420 BP.

XX AC AAX35711;

XX DT 09-JUL-1999 (first entry)

XX DE cDNA encoding a protein identified by the signal sequence trap method.

XX KW Signal sequence trap method; SST method; immunisation; inhibition;
 KW infection; allergy; cancer; regulation; tissue formation; tissue repair;
 KW activin activity; inhibin activity; chemokine activity;
 KW cytokine activity; blood coagulation regulation; agonist; antagonist;
 KW metabolic disorder; hormonal disorder; immune disorder;
 KW severe combined immunodeficiency; SCID; AIDS; thrombosis; cancer;
 XX wound; ss.

XX OS Homo sapiens.

XX PN W09918126-A1.

XX PD 15-APR-1999.

XX PF 06-OCT-1998; 98WO-JP04514.

XX PR 07-OCT-1997; 97JP-0274674.

XX PA (ONOY) ONO PHARM CO LTD.

XX PI Fukushima D, Shibayama S, Tada H;